

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2005, 10:05:19 ; Search time 169 Seconds

(without alignments)
2213.002 Million cell updates/sec

Title: US-09-989-687-126

Perfect score: 5287

Sequence: 1 MRAVPEGFRKRLGSDMGN.....CDPLKKPKHFIDFTMAECS 967

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04.*

1: Geneseqp1980s.*

2: Geneseqp1990s.*

3: Geneseqp2000s.*

4: Geneseqp2001s.*

5: Geneseqp2002s.*

6: Geneseqp2003as.*

7: Geneseqp2003bs.*

8: Geneseqp2004s.*

* Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5232	99.0	967	2	AAW80285 Human int
2	5224	98.8	967	2	AAW80285 Human Tan
3	5224	98.8	967	8	ADR14133 Human NP-
4	5224	98.8	967	8	ADR14133 Human NP-
5	5224	98.8	967	8	ADQ39940 Human myo
6	5221.5	98.8	968	4	ADQ39941 Human myo
7	5201	98.4	967	2	AAW80285 Human sec
8	5201	98.4	967	2	AAW80285 Human sec
9	5201	98.4	967	6	ADA41003 Human sec
10	5201	98.4	967	7	ADB91631 Human sec
11	5201	98.4	967	7	ADB91631 Human sec
12	5201	98.4	967	7	ADQ37948 Human sec
13	5144	97.3	950	2	AAW80285 Human MET
14	5144	97.3	950	4	AAW80285 Human ADA
15	5144	97.3	950	4	AAW80285 Human MET
16	5140	97.2	949	7	ADG72483 Human agg
17	5136	97.1	950	8	ADQ20218 Human PRO
18	5136	97.1	950	8	ADQ39942 Human PRO
19	4293.5	79.7	968	6	ABU08387 Murine ma
20	4215.5	79.7	950	3	AAW80285 Amino aci
21	3996.5	75.6	896	3	AAW80285 Mouse met
22	3922	74.2	727	2	AAW80285 Human ADA
23	2495.5	47.2	924	5	ABP70062 Human NOV
24	2477.5	46.9	950	5	AAW80285 Human pro
25	2477.5	46.9	950	6	ABU09520 Human pro

26	2477.5	46.9	950	8	ADQ88214 Human 655
27	2473.5	46.8	950	6	ABR40092 Human ADA
28	2471.5	46.7	950	4	AAW80285 Human met
29	2455.5	46.4	952	5	AAU74751 Human pro
30	2408	45.5	928	5	AAU72899 Human met
31	2323.5	43.9	890	6	ABP96306 Human ADA
32	2321.5	43.9	890	8	ADK70513 Respirato
33	2320.5	43.9	890	2	AAW80285 Human MET
34	2320.5	43.9	890	4	AAW80285 Human MET
35	2313	43.7	889	4	AAW80285 Human ADA
36	2265.5	42.9	905	4	AAW80285 Murine AD
37	2190	41.4	823	6	ABU08383 Human mat
38	2129	40.3	837	4	AAW80285 Human egg
39	2129	40.3	837	7	ADB85488 Human egg
40	2129	40.3	840	3	AAW80285 Human met
41	2124	40.2	837	2	AAW80285 Human egg
42	2124	40.2	837	7	ADJ69542 Human hea
43	2124	40.2	837	8	ADS20209 Human egg
44	2121	40.1	846	8	ADS20232 Human egg
45	2117	40.0	837	3	AAW80285 Human PRO

ALIGNMENTS

RESULT 1
AAW80285
ID AAW80285 standard; protein; 967 AA.
XX
AC AAW80285;
XX
DT 19-JAN-1999 (first entry)
XX
DE Human integrin ligand polypeptide ITGL-TSP.
XX

KW ITGL-TSP; integrin ligand; angiogenic disease; cancer; atherosclerosis;
KW Chronic inflammatory disorder; rheumatoid arthritis; tissue remodeling;
KW macular degeneration; diabetic retinopathy; Alzheimer's disease; human;
KW restenosis.

OS Homo sapiens.

XX EP874050-A2.

XX 28-OCT-1998.

XX 27-JAN-1998; 98EP-00300575.

XX 24-APR-1997; 97US-00845496.

XX (SMIK) SMITHKLINE BEECHAM CORP.

XX (SMIK) SMITHKLINE BEECHAM PLC.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Jonak ZL, Trulli SH, Fronwald JA, Hastings GA, Terrett JA;

XX WPI; 1998-544643/47.

XX N-PSDB; AAW66508.

XX DNA encoding integrin ligand polypeptide ITGL-TSP - used to treat

XX angiogenic diseases, restenosis, Alzheimer's disease and in tissue

XX remodeling.

XX Claim 11; Page 6-9; 24pp; English.

XX This represents a human integrin ligand polypeptide ITGL-TSP. ITGL-TSP

XX polypeptides can be used in the treatment of angiogenic diseases such as

XX cancer, cancer metastasis, chronic inflammatory disorders, rheumatoid

XX arthritis, atherosclerosis, macular degeneration or diabetic retinopathy,

XX restenosis, Alzheimer's disease and tissue remodeling. They can be used

XX to treat a subject in need of enhanced activity or expression of the ITGL

XX -TSP polypeptide

SQ	Sequence 967 AA;	
Query Match	99.0%; Score 5232; DB 2; Length 967;	
Best Local Similarity	99.0%; Pred. No. 0;	
Matches	957; Conservative 4; Mismatches 6; Indels 0; Gaps 0;	
Qy	1 MORAVPEGFRKRLGSDMGNAERAPGSRSGFVPPTLLLLAAALLAVSDALGRPSEDEEL 60	
Db	1 MORAVPEGFRKRLGSDMGNAERAPGSRSGFVPPTLLLLAAALLAVSDALGRPSEDEEL 60	
Qy	61 VVPELERVPGHGTTRLRHLHAFDQQLDLVDPDSFLAPGFTLQNVGRKSGSDTLPETDL 120	
Db	61 VVPELERAPGHGTTRLRHLHAFDQQLDLVDPDSFLAPGFTLQNVGRKSGSETLPETDL 120	
Qy	121 AHCFYSGTVNGDPSAAALSCEGVRGAFYLLGEAYFIQPLPAASERLATAAPEKPPAP 180	
Db	121 AHCFYSGTVNGDPSAAALSCEGVRGAFYLLGEAYFIQPLPAASERLATAAPEKPPAP 180	
Qy	181 LQFHLLRRNRQDVGCTCGVVDDEPRPTGKAETEDDEGTEGEGPQWSPQDPALQGVG 240	
Db	181 LQFHLLRRNRQDVGCTCGVVDDEPRPTGKAETEDDEGTEGEGPQWSPQDPALQGVG 240	
Qy	241 OPTGTGSIKRRFVSSHRYVETMLVADQSMAEFHGSLKHYLLTLFSAARLYKHPISRN 300	
Db	241 OPTGTGSIKRRFVSSHRYVETMLVADQSMAEFHGSLKHYLLTLFSAARLYKHPISRN 300	
Qy	301 SVSLVWVKILVIHDEQKPEVTSNAALTFLRNCWKQHNPPSRDABHYDTAILFTROD 360	
Db	301 SVSLVWVKILVIHDEQKPEVTSNAALTFLRNCWKQHNPPSRDABHYDTAILFTROD 360	
Qy	361 LCGSQTCDTLGADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFMNPHDDAKQACSL 420	
Db	361 LCGSQTCDTLGADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFMNPHDDAKQACSL 420	
Qy	421 NGVNQDSHMASLNLNLDHSPFSPCSGYMTSFLDNGHGCLMDKQNPLOLPDLPGT 480	
Db	421 NGVNQDSHMASLNLNLDHSPFSPCSGYMTSFLDNGHGCLMDKQNPLOLPDLPGT 480	
Qy	481 SYDANROCOFTFGEDSKHCPDAASTCTLWCTGTSGGVILVCOTKHPFWDGTSCEGKWC 540	
Db	481 SYDANROCOFTFGEDSKHCPDAASTCTLWCTGTSGGVILVCOTKHPFWDGTSCEGKWC 540	
Qy	541 INGCNVNKHNRKHDPDPFHGSGWGMWPGDCSRCTCGGVQVYTMRECDNPVPKNGKCEG 600	
Db	541 INGCNVNKHNRKHDPDPFHGSGWGMWPGDCSRCTCGGVQVYTMRECDNPVPKNGKCEG 600	
Qy	601 KRVYRSCNLEDCPDNNGKTFREOCEAHNEFSAKSGGPAVEWIPKYAGVSPKDRCKL 660	
Db	601 KRVYRSCNLEDCPDNNGKTFREOCEAHNEFSAKSGGPAVEWIPKYAGVSPKDRCKL 660	
Qy	661 ICQAKGIGYFVLQPKVVDGTPCSPDSTSVQVQGVKAGCDRIIDSKKFKDCKGVCGGN 720	
Db	661 ICQAKGIGYFVLQPKVVDGTPCSPDSTSVQVQGVKAGCDRIIDSKKFKDCKGVCGGN 720	
Qy	721 GSTCKKISGVSATKPGVHDIITPTGATNIEVKORNGSRNNGSFLAIKAAAGTIVLN 780	
Db	721 GSTCKKISGVSATKPGVHDIITPTGATNIEVKORNGSRNNGSFLAIKAAAGTIVLN 780	
Qy	781 GDYTLSTLEQDQIMYKGVVLYRSGSSAALERIRSFSPLEKIPTIOVLTVGNALRPKIYTY 840	
Db	781 GDYTLSTLEQDQIMYKGVVLYRSGSSAALERIRSFSPLEKIPTIOVLTVGNALRPKIYTY 840	
Qy	841 FVKKKESFNAIPTFSAWVIEWECSEKSCBLGHORRLVECRDINGQPASECAKEVPAS 900	
Db	841 FVKKKESFNAIPTFSAWVIEWECSEKSCBLGHORRLVECRDINGQPASECAKEVPAS 900	
Qy	901 TRPCADHPCPQWQLGEWSSCKTCGKYKTSKLCSLHDGVLGSHDSCDPLKPKHFTDF 960	
Db	901 TRPCADHPCPQWQLGEWSSCKTCGKYKTSKLCSLHDGVLGSHDSCDPLKPKHFTDF 960	
Qy	961 CTMAECS 967	
Db	961 CTMAECS 967	

RESULT 2	
AA04142	
ID	AA04142 standard; protein; 967 AA.
XX	AA04142;
XX	15-JUN-1999 (first entry)
XX	Human Tango-71 protein.
XX	Human; Tango-71; Tango-73; Tango-74; Tango-76; Tango-83; diagnosis;
XX	detection.
XX	Homo sapiens.
XX	WO9907850-A1.
XX	18-FEB-1999.
XX	06-AUG-1998; 98WO-US016502.
XX	06-AUG-1997; 97US-0054966P.
XX	05-SEP-1997; 97US-0058108P.
XX	(MILL-) MILLENNIUM BIOTHERAPEUTICS INC.
XX	Holtzman DA, Goodearl ADJ;
XX	WPI; 1999-167426/14.
XX	N-PSDB; AAX19955.
XX	New TANGO polypeptides and nucleic acids encoding them - useful as
XX	diagnostic agents and for treating disorders caused by aberrant
XX	expression of TANGO.
XX	Claim 8; Fig 1; 84pp; English.
XX	The present sequence represents human Tango-71. Tango polypeptides are
XX	useful for identifying compounds which bind the polypeptide via direct
XX	binding, competition binding assays or Tango-71, -73, -74, 76 or -83-
XX	mediated signal transduction. Tango polypeptides are also useful for
XX	identifying modulating compounds by determining effect on Tango activity.
XX	Tango polypeptides and nucleic acids are useful for diagnosing diseases
XX	related to aberrant expression of Tango, and Tango polypeptides are
XX	useful for raising antibodies which can be used in diagnostic assays for
XX	detection of Tango, and also for generating anti-idiotypic antibodies for
XX	prevention and protection
SQ	Sequence 967 AA;
Query Match	98.8%; Score 5224; DB 2; Length 967;
Best Local Similarity	98.9%; Pred. No. 0;
Matches	956; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
Qy	1 MORAVPEGFRKRLGSDMGNAERAPGSRSGFVPPTLLLLAAALLAVSDALGRPSEDEEL 60
Db	1 MORAVPEGFRKRLGSDMGNAERAPGSRSGFVPPTLLLLAAALLAVSDALGRPSEDEEL 60
Qy	61 VVPELERVPGHGTTRLRHLHAFDQQLDLVDPDSFLAPGFTLQNVGRKSGSDTLPETDL 120
Db	61 VVPELERAPGHGTTRLRHLHAFDQQLDLVDPDSFLAPGFTLQNVGRKSGSETLPETDL 120
Qy	121 AHCFYSGTVNGDPSAAALSCEGVRGAFYLLGEAYFIQPLPAASERLATAAPEKPPAP 180
Db	121 AHCFYSGTVNGDPSAAALSCEGVRGAFYLLGEAYFIQPLPAASERLATAAPEKPPAP 180
Qy	181 LQFHLLRRNRQDVGCTCGVVDDEPRPTGKAETEDDEGTEGEGPQWSPQDPALQGVG 240
Db	181 LQFHLLRRNRQDVGCTCGVVDDEPRPTGKAETEDDEGTEGEGPQWSPQDPALQGVG 240
Qy	241 OPTGTGSIKRRFVSSHRYVETMLVADQSMAEFHGSLKHYLLTLFSAARLYKHPISRN 300

Db 241 QPTGTSIRKRFVSHRYVETMLVADQSMAEFHGSLKHYLLTLTFSVAARLYKPEIRN 300
Qy 301 SVSLVVKLVLDHDKQKPEVTSNAALTLRNFQKQKQHPSPDRDAEHYDTAILFTTQD 360
Db 301 SVSLVVKLVLDHDKQKPEVTSNAALTLRNFQKQKQHPSPDRDAEHYDTAILFTTQD 360
Qy 361 LCGSQTCDTLGMADVGTVCDSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQOCASL 420
Db 361 LCGSQTCDTLGMADVGTVCDSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQOCASL 420
Qy 421 NGVNDSHMMASMLNLDHSPQWSPSCSGYMITSLDNGHGECIMDKPQNPQIQLPGDLPGT 480
Db 421 NGVNDSHMMASMLNLDHSPQWSPSCSGYMITSLDNGHGECIMDKPQNPQIQLPGDLPGT 480
Qy 481 SYDANRQCOFTGEGDSKCPDAASTCTLWCTGTSGVLVLCOTKHPFPAWADGTSCEGKWC 540
Db 481 SYDANRQCOFTGEGDSKCPDAASTCTLWCTGTSGVLVLCOTKHPFPAWADGTSCEGKWC 540
Qy 541 INGKCVNKHKHFDTPFHSGWGMGPMGDCSRTCGGVQVYTMRECDNPVPKNGGKYCEG 600
Db 541 INGKCVNKHKHFDTPFHSGWGMGPMGDCSRTCGGVQVYTMRECDNPVPKNGGKYCEG 600
Qy 601 KVRVRSNLEDCPDNNGKTFREOCEAHNEFSKASFGSGPAVEWIPKYAGVSPKORCKL 660
Db 601 KVRVRSNLEDCPDNNGKTFREOCEAHNEFSKASFGSGPAVEWIPKYAGVSPKORCKL 660
Qy 661 ICQAKGIGVFFVLQPKVVDGTPCSPDSTSVQGVQKAGCDRIIDSKKPKDCKGCGGN 720
Db 661 ICQAKGIGVFFVLQPKVVDGTPCSPDSTSVQGVQKAGCDRIIDSKKPKDCKGCGGN 720
Qy 721 GSTCKKISGSVTSAPGYHDIITPTGATNIEVKQNRQSRNNGSFLAIKAADGTIYL 780
Db 721 GSTCKKISGSVTSAPGYHDIITPTGATNIEVKQNRQSRNNGSFLAIKAADGTIYL 780
Qy 781 GDYTLSTLEQDIWYKGVVLRYSGSSAALIERISFPLKEPLTIQVLTGNALRPKIKTY 840
Db 781 GDYTLSTLEQDIWYKGVVLRYSGSSAALIERISFPLKEPLTIQVLTGNALRPKIKTY 840
Qy 841 FVKKKESFNALPTFSAMVIBEWGCSKSCELGWORRLVECDINGQPASECAKEVKPAS 900
Db 841 FVKKKESFNALPTFSAMVIBEWGCSKSCELGWORRLVECDINGQPASECAKEVKPAS 900
Qy 901 TRPCADHPCPQWQLGEWSKSCSTCGKYKTKSLKLSHDGGVLSHSDCDPLKKPKHFIDF 960
Db 901 TRPCADHPCPQWQLGEWSKSCSTCGKYKTKSLKLSHDGGVLSHSDCDPLKKPKHFIDF 960
Qy 961 CTMAECS 967
Db 961 CTMAECS 967
RESULT 3
ADRI14133
ID ADRI14133 standard; protein; 967 AA.
AC ADRI14133;
XX
DT 21-OCT-2004 (first entry)
DE Human NF-kappaB pathway-associated protein SeqID134.
KW NF-kappaB pathway; antiinflammatory; cytostatic; hepatotropic; virucide;
KW antiarthritic; antiarthritis; gastrointestinal-Gen; antiarthritic;
KW antiarteriosclerotic; immunomodulator; cerebroprotective; vasotropic;
KW immunosuppressive; vulnary; gene therapy; immune disorder;
KW inflammatory disorder; NF-kappaB regulation; cancer; aberrant
KW hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour;
KW hyper-IgM syndrome; hypohidrotic ectodermal dysplasia;
KW X-linked anhidrotic ectodermal dysplasia; immunodeficiency;
KW viral infection; HIV-1; HTLV-1; hepatitis B; hepatitis C; EBV; influenza;
KW viral replication; host cell survival; evasion of immune response;
KW rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;

KW atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;
KW autoimmune disorder; hyper immune activity;
KW aberrant acute phase response; hypercongenital condition; birth defect;
KW necrotic lesion; wound; organ transplant rejection;
KW aberrant signal transduction; proliferating disorder; cancer;
KW HIV propagation; human.
XX
OS Homo sapiens.
XX
PN WO2004065577-A2.
XX
PD 05-AUG-2004.
PF 13-JAN-2004; 2004WO-US000798.
PR 14-JAN-2003; 2003US-0440068P.
PR 12-MAY-2003; 2003US-0469757P.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Nadler SG, Neubauer MG, Feder JN, Carman J;
XX
DR WPI: 2004-562168/54.
DR N-PSDB; ADRI14132.
XX
PT New isolated polynucleotides and polypeptides associated with NF-kappaB
PT pathway, useful for diagnosing, treating, or preventing disorders or
PT diseases associated with NF-kappaB pathway.
XX
PS Claim 6; SEQ ID NO 134; 237pp; English.
XX
CC This invention relates to the novel association of protein sequences (and
CC the genes which encode them) to the NF-kappaB pathway. The invention may
CC be useful for the production of compounds with an antiinflammatory,
CC cyostatic, hepatotropic, virucide, antiarthritic, antiarthritis,
CC gastrointestinal-Gen, antiasthmatic, antiarteriosclerotic,
CC immunomodulator, cerebroprotective, vasotropic, immunosuppressive or
CC vulnary activity or for gene therapy. The proteins and nucleotides are
CC useful for diagnosing, preventing, treating, or ameliorating conditions
CC or diseases associated with the NF-kappaB pathway. The condition is an
CC immune disorder, an inflammatory disorder, an inflammatory disorder
CC related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis,
CC hepatic disorders, Hodgkin's lymphomas, haematopoietic tumours, hyper-IgM
CC syndromes, hypohidrotic ectodermal dysplasia, X-linked anhidrotic
CC ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, HTLV-1,
CC hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell
CC survival, evasion of immune responses, rheumatoid arthritis, inflammatory
CC bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick
CC syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper
CC immune activity, disorders related to aberrant acute phase responses,
CC hypercongenital conditions, birth defects, necrotic lesions, wounds,
CC organ transplant rejection, conditions related to organ transplant
CC rejection, disorders related to aberrant signal transduction.
CC proliferating disorders, cancers and HIV propagation in cells infected
CC with other viruses. The present sequence is that of a human protein which
CC is subject to the novel association with the NF-kappaB pathway of the
CC invention. Note: This sequence does not appear in the specification but
CC was obtained by the indexer from Genbank.
XX
SQ Sequence 967 AA;

Query Match 98.8%; Score 5224; DB 8; Length 967;
Best Local Similarity 98.8%; Pred. NO. 0;
Matches 955; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Qy 1 MQRVPEFGFRKLGSDMGNAERAPGSGFVPVTLTLLAALLAVSDALGPSESEEL 60
Db 1 MQRVPEFGFRKLGSDMGNAERAPGSGFVPVTLTLLAALLAVSDALGPSESEEL 60
Qy 61 VVPELERYPGHGTTRRLRHLHAFDQDLDPVPPDSSFLAPGFTLQNVGRKSGSDTLPETDL 120
Db 61 VVPELERYPGHGTTRRLRHLHAFDQDLDPVPPDSSFLAPGFTLQNVGRKSGSDTLPETDL 120

Qy	121	AHCFYSGTVNGDPSSAALSLCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGEKPPAP	180
Db	121		
Qy	181	LQFHLLRRNRQGVGTCGVVDDEPRPTKCAFTEDDEGTEDEGEGPQWSPQDPALQGVG	240
Db	181		
Qy	241	QPTGTGSTRKRKRFVSSHRYVETMLVADQSMABFHGSGLKHVLLTLFVSVAARLYKHPSIRN	300
Db	241		
Qy	301	SVSLVVVKILVTHDQKGPVTSNAALTLRNFCKWQKOHNPSPSDRDAEHYDTAILFTQD	360
Db	301		
Qy	361	LCGSQTCDTLGMADVTCDFPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQKQASL	420
Db	361		
Qy	421	NGVNODSHMAASMLNLDHSPQWSPCSGYMITSFLDNHGHECLMDKPNPTQLPGDLPGT	480
Db	421		
Qy	481	SYDANRQCOFTFGEDSKHCPDAASTCTLWCTGTSGGVLVCQTKHFPWADGTSCEGKWC	540
Db	481		
Qy	541	INGKCVNKHRRKHFDTPPHGSGMWGPGDCSRTCCTGGGVQVYTMRECDNPPVKNKGKCEG	600
Db	541		
Qy	601	KRVYRSCNLEDCPDNNKCTFREOCEAHNFEKASFGSGPAVEWIPKYAGVSPKDRCKL	660
D5	601		
Qy	661	ICOAKGIGVFFVLPQKVVDGTPCSPDSTSVCGQCVKAGCDRIIDSKKFPDKGVCVCGN	720
Db	661		
Qy	721	GSTCKKISGVSATKPGYHDIITPTGATNIEVKORNGSRNNGSFLAIKAADGTYILN	780
Db	721		
Qy	781	GDYTLSTLEQDIMYKGVVLRVSGSSAALERIRSFSPKLEPTIOVLTVGNALRPKIKYTY	840
Db	781		
Qy	841	FVKKKESFNAIPTFSAMVIEWGECSKSCBLGWORLVECRDINGQPASECAKEVPKAS	900
Db	841		
Qy	901	TRPCADHPCPQWLGEWSSCKTCGKYKTSLSKLSHGDGVLGSHDSCDPLKPKGHFIDF	960
Db	901		
Qy	961	CTMAECS 967	
Db	961		

RESULT 4

ID	ADQ39940	standard; protein; 967 AA.
XX		
AC	ADQ39940;	
XX		
DT	18-NOV-2004	(first entry)
XX		
DE	Human myocardial infarction-associated gene derived protein, SEQ ID 1603.	
XX		
KW	Myocardial infarction; detection; single nucleotide polymorphism; SNP;	
KW	cardiant; gene therapy; human.	

XX		Homo sapiens.	
OS			
XX		WO2004058052-A2.	
FN			
XX		15-JUL-2004.	
PD			
XX			
XX		22-DEC-2003; 2003WO-US040978.	
PF			
XX		20-DEC-2002; 2002US-0434778P.	
PR		10-MAR-2003; 2003US-0453135P.	
PR		30-APR-2003; 2003US-0466412P.	
PR		23-SEP-2003; 2003US-0504955P.	
XX			
FA		(APPL-) APPLERA CORP.	
XX			
PI		Cargill M, Devlin JJ, Iakoubova O;	
XX			
XX		WPI; 2004-533949/51.	
DR		N-PSDB; ADQ39112.	
XX			
PT		Identifying an individual who has an altered risk for developing	
PT		myocardial infarction by detecting a single nucleotide polymorphism in	
PT		the individual's nucleic acids.	
XX		Claim 10; SEQ ID NO 1603; 145pp; English.	
PS			
XX		The invention relates to a novel method for identifying an individual who	
CC		has an altered risk for developing myocardial infarction. The method	
CC		comprises detecting a single nucleotide polymorphism (SNP) in any one of	
CC		the nucleotide sequences given in the specification in the individual's	
CC		nucleic acids, where the presence of the SNP is correlated with an	
CC		altered risk for myocardial infarction in the individual. The invention	
CC		further comprises: an isolated nucleic acid molecule comprising at least	
CC		8 contiguous nucleotides where one of the nucleotides is an SNP given in	
CC		the specification or its complement and encoding any one of the amino	
CC		acid sequences given in the specification; an isolated polypeptide	
CC		comprising an amino acid sequence given in the specification; an antibody	
CC		that specifically binds to the polypeptide or its antigen-binding	
CC		fragment; an amplified polynucleotide containing an SNP given in the	
CC		specification and which is between about 16 and 1000 nucleotides in	
CC		length; a kit for detecting an SNP in a nucleic acid, comprising the	
CC		polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a	
CC		nucleic acid molecule; a method of detecting a variant polypeptide; and a	
CC		method for identifying an agent useful in treating or preventing	
CC		myocardial infarction. The novel detection method has cardiac activity.	
CC		The nucleic acids of the invention may be used in gene therapy. The	
CC		method is useful in identifying an individual who has an increased or	
CC		decreased risk for developing myocardial infarction and for preparing a	
CC		composition for treating or preventing myocardial infarction. This	
CC		sequence represents the protein of a human myocardial infarction-	
CC		associated gene containing one or more SNP's of the invention. Note: This	
CC		CC sequence not shown in the specification. The sequence has come from	
CC		an electronic sequence listing downloaded from the WIPO website.	
XX			
SQ		Sequence 967 AA;	

Query Match	98.8%;	Score 5224;	DB 8;	Length 967;
Best Local Similarity	98.9%;	Pred. No. 0;		
Matches 956;	Conservative 4;	Mismatches 7;	Indels 0;	Gaps 0;

Qy	1	MQRVPEGFGRRKLGSDMGNAERAPGSRSGPVPPTLLLLAAALIAVSDALGRPSEDEEL	60
Db	1	MQRVPEGFGRRKLGSDMGNAERAPGSRSGPVPPTLLLLAAALIAVSDALGRPSEDEEL	60
Qy	61	VVPELERVPCGGTTRLRHLHAFDQQLDLDVPPDSFLAPGFTLQNVGRKSGSDTLPETDL	120
Db	61	VVPELERAPCGTTRLRHLHAFDQQLDLELRPDSFLAPGFTLQNVGRKSGSETPLPETDL	120
Qy	121	AHCFYSGTVNGDPSSAALSLCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGEKPPAP	180
Db	121	AHCFYSGTVNGDPSSAALSLCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGEKPPAP	180

```
QY 181 LQFHLRRNRQGVGGTCTGVDDRRPTGKARTEDDEGTEDEGSGPQSPDPAALQGVG 240
Db 181 LQFHLRRNRQGVGGTCTGVDDRRPTGKARTEDDEGTEDEGSGPQSPDPAALQGVG 240
QY 241 OPTGTGSIKKEFVSHRVETMLVADQSMAEFHGSLKHYLLTLFSAARLYKHPISRN 300
Db 241 OPTGTGSIKKEFVSHRVETMLVADQSMAEFHGSLKHYLLTLFSAARLYKHPISRN 300
QY 301 SVSLVVVKILVHDSKQPEVTSNAALTILRNFENQKQHNPPSDRDAEHYDTAILFTTQD 360
Db 301 SVSLVVVKILVHDSKQPEVTSNAALTILRNFENQKQHNPPSDRDAEHYDTAILFTTQD 360
QY 361 LCGSQTCDTLGNADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDADAKQASL 420
Db 361 LCGSQTCDTLGNADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDADAKQASL 420
QY 421 NGVNDSHMMASMLNLDHSDQSPSCSGYMITSLFDNGHGECLMDKPNPTQLPGDLPGT 480
Db 421 NGVNDSHMMASMLNLDHSDQSPSCSGYMITSLFDNGHGECLMDKPNPTQLPGDLPGT 480
QY 481 SYDANRQCOFTFGEDSKHCPDAASTCSTLWCTGTSGGVLVCOQTHFPWADGTSCGEGKWC 540
Db 481 SYDANRQCOFTFGEDSKHCPDAASTCSTLWCTGTSGGVLVCOQTHFPWADGTSCGEGKWC 540
QY 541 INKCVNKNRKHFTPTPHGSGWMPGDCSRTCCGGVQYTMRECDNPVPKNGKCYCEG 600
Db 541 INKCVNKTDRKHFTPTPHGSGWMPGDCSRTCCGGVQYTMRECDNPVPKNGKCYCEG 600
QY 601 KRVVRSCNLEDCPNNGKTFREOCEAHNPFKSGFSGPAVEWIPKYAGVSPKDRCKL 660
Db 601 KRVVRSCNLEDCPNNGKTFREOCEAHNPFKSGFSGPAVEWIPKYAGVSPKDRCKL 660
QY 661 ICOAKGIGYFFVLQPKVVDGTPCSPDSTSVQGVQCVKAGCDRIIDSKKKFKDKGVCNGN 720
Db 661 ICOAKGIGYFFVLQPKVVDGTPCSPDSTSVQGVQCVKAGCDRIIDSKKKFKDKGVCNGN 720
QY 721 GSTCKKISGVSATKPGVHDIIITPTGATNIEVKORNGSRNNGSFLAIAKADGTYILN 780
Db 721 GSTCKKISGVSATKPGVHDIIITPTGATNIEVKORNGSRNNGSFLAIAKADGTYILN 780
QY 781 GDTLTSTLEQDLMYKGVVLRYSGSSAALERIRSFSPLEPLTIQVLTGNALRPKIKTY 840
Db 781 GDTLTSTLEQDLMYKGVVLRYSGSSAALERIRSFSPLEPLTIQVLTGNALRPKIKTY 840
QY 841 FVKKKESFNALPTFSANVIEBWGSCSKSCELGMQRRLVECRDINGQPASECAKEVKPAS 900
Db 841 FVKKKESFNALPTFSANVIEBWGSCSKSCELGMQRRLVECRDINGQPASECAKEVKPAS 900
QY 901 TRPCADHPCPQWLQGEWSSCKTKCGYKKTSLKCLSHDGGVLSHSDCPDLKKPKHFTDF 960
Db 901 TRPCADHPCPQWLQGEWSSCKTKCGYKKTSLKCLSHDGGVLSHSDCPDLKKPKHFTDF 960
QY 961 CTMAECS 967
Db 961 CTMAECS 967
RESULT 5
ADQ39941
ID ADQ39941 standard; protein; 967 AA.
XX
AC
XX
XX
DT 18-NOV-2004 (first entry)
XX
DE Human myocardial infarction-associated gene derived protein, SEQ ID 1604.
XX
KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;
KW cardiac; gene therapy; human.
XX
OS Homo sapiens.
XX
PN WO2004058052-A2.
```

```
XX 15-JUL-2004.
PD
XX
PF 22-DEC-2003; 2003WO-US040978.
XX
PR 20-DEC-2002; 2002US-0434778P.
PR 10-MAR-2003; 2003US-0453135P.
PR 30-APR-2003; 2003US-0466412P.
PR 23-SEP-2003; 2003US-0504955P.
XX
PA (APPL-) APPLERA CORP.
XX
PI Cargill M, Devlin JJ, Iakoubova O;
XX
DR WPI; 2004-533949/51.
DR N-PSDB; ADQ39113.
XX
PT Identifying an individual who has an altered risk for developing
PT myocardial infarction by detecting a single nucleotide polymorphism in
PT the individual's nucleic acids.
XX
PS Claim 10; SEQ ID NO 1604; 145pp; English.
XX
CC The invention relates to a novel method for identifying an individual who
CC has an altered risk for developing myocardial infarction. The method
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
CC the nucleotide sequences given in the specification in the individual's
CC nucleic acids, where the presence of the SNP is correlated with an
CC altered risk for myocardial infarction in the individual. The invention
CC further comprises: an isolated nucleic acid molecule comprising at least
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
CC the specification or its complement and encoding any one of the amino
CC acid sequences given in the specification; an isolated polypeptide
CC comprising an amino acid sequence given in the specification; an antibody
CC that specifically binds to the polypeptide or its antigen-binding
CC fragment; an amplified polynucleotide containing an SNP given in the
CC specification and which is between about 16 and 1000 nucleotides in
CC length; a kit for detecting an SNP in a nucleic acid, comprising the
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
CC method for identifying an agent useful in treating or preventing
CC myocardial infarction. The novel detection method has cardiac activity.
CC The nucleic acids of the invention may be used in gene therapy. The
CC method is useful in identifying an individual who has an increased or
CC decreased risk for developing myocardial infarction and for preparing a
CC composition for treating or preventing myocardial infarction. This
CC sequence represents the protein of a human myocardial infarction-
CC associated gene containing one or more SNPs of the invention. Note: This
CC sequence was not shown in the specification. The sequence has come from
CC an electronic sequence listing downloaded from the WIPO website.
XX
SQ Sequence 967 AA;
Query Match 98.8%; Score 5224; DB 8; Length 967;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 956; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
QY 1 MORAVPGEFGRKLGSDMGNAERAPGSRSGFVPTLLLLAALLAVSDALGRPSDEEL 60
Db 1 MORAVPGEFGRKLGSDMGNAERAPGSRSGFVPTLLLLAALLAVSDALGRPSDEEL 60
QY 61 VVPELERVPGHGTTRLRHAFDQQLDLDVPPDSSFLAPGFTLQNVGRKSGSDTLPETDL 120
Db 61 VVPELERVPGHGTTRLRHAFDQQLDLDVPPDSSFLAPGFTLQNVGRKSGSDTLPETDL 120
QY 121 AHCFYSGTVNGDPSAAALSLCEGVGAFYLLGAYFIQPLPAASERLATAPAEKPPAP 180
Db 121 AHCFYSGTVNGDPSAAALSLCEGVGAFYLLGAYFIQPLPAASERLATAPAEKPPAP 180
QY 181 LQFHLRRNRQGVGGTCTGVDDRRPTGKARTEDDEGTEDEGSGPQSPDPAALQGVG 240
Db 181 LQFHLRRNRQGVGGTCTGVDDRRPTGKARTEDDEGTEDEGSGPQSPDPAALQGVG 240
```


Db 181 PLQFHLRRNRQDVGTCGVVDDEPRPTGKAEDEDEBGTGEDEGPQWSPQDPALQGV 240
QY 240 GOPTGTGSRKRRFVSSHRYVETMLVADQSMAEFHGSLGKHYLLTLFSAARLYKHPsir 299
Db 241 GOPTGTGSRKRRFVSSHRYVETMLVADQSMAEFHGSLGKHYLLTLFSAARLYKHPsir 300
QY 300 NSVSLVWVKILVIHDEQKPEVTSNAALTNRNFCWQKQHNPPSDRDAEHYDTAILPFRQ 359
Db 301 NSVSLVWVKILVIHDEQKPEVTSNAALTNRNFCWQKQHNPPSDRDAEHYDTAILPFRQ 360
QY 360 DLGSGSOTCDTLGMADVGTCDPSRSCSVIEDGLQAAFTTAHELGHVFNMPHDDAKQAS 419
Db 361 DLGSGSOTCDTLGMADVGTCDPSRSCSVIEDGLQAAFTTAHELGHVFNMPHDDAKQAS 420
QY 420 LNGVNQDSHMASMLNLDHSPWSPSCGYMITSLDNGHGCECLMDKPNQPIQLPGLDLP 479
Db 421 LNGVNQDSHMASMLNLDHSPWSPSCGYMITSLDNGHGCECLMDKPNQPIQLPGLDLP 480
QY 480 TSYDANROCOFTFGSDSKHCPDAACTSLMCTGTSGGVLVCQTKHFPWADTSCGEGKW 539
Db 481 TSYDANROCOFTFGSDSKHCPDAACTSLMCTGTSGGVLVCQTKHFPWADTSCGEGKW 540
QY 540 CINGKCVNKNRKHFDTPPHGSGMWGPGDCSRTC GGSGVQVTMBECNDPVPKNGKYCE 599
Db 541 CINGKCVNKNRKHFDTPPHGSGMWGPGDCSRTC GGSGVQVTMBECNDPVPKNGKYCE 600
QY 600 GKRVRYSCLNEDCPDNNKGTFRBQCEAHNEFSAFSGSPAVIEWIPKYAGVSPKDRCK 659
Db 601 GKRVRYSCLNEDCPDNNKGTFRBQCEAHNEFSAFSGSPAVIEWIPKYAGVSPKDRCK 660
QY 660 LICQAKGIGYFVLQPKVVDGTPCSPDSTSCVQGCVKAGCDRIIDSKKPKDKCGVCGG 719
Db 661 LICQAKGIGYFVLQPKVVDGTPCSPDSTSCVQGCVKAGCDRIIDSKKPKDKCGVCGG 720
QY 720 NGSTCKKISGVSYSKAPGVHDIITPTGATNIEVKQRNQRNGSFLAIKAAADGTYL 779
Db 721 NGSTCKKISGVSYSKAPGVHDIITPTGATNIEVKQRNQRNGSFLAIKAAADGTYL 780
QY 780 NGDYTLSTLEQDQIMYGVVLRYSGSSAALERSFSPKLEPTIQVLTGVNLRPKIKYT 839
Db 781 NGDYTLSTLEQDQIMYGVVLRYSGSSAALERSFSPKLEPTIQVLTGVNLRPKIKYT 840
QY 840 YFVKKKESFNAIPFTSAWVIEWEGECSKCSBLGWQRRLVECRDINGQPASECAKEVKA 899
Db 841 YFVKKKESFNAIPFTSAWVIEWEGECSKCSBLGWQRRLVECRDINGQPASECAKEVKA 900
QY 900 STRPCADHPCPQWOLGWSKSKTCGKYKTSKLCLSHDGGVLSHDCDPLKPKHFD 959
Db 901 STRPCADHPCPQWOLGWSKSKTCGKYKTSKLCLSHDGGVLSHDCDPLKPKHFD 960
QY 960 FCTMAECS 967
Db 961 FCTMAECS 968

RESULT 7

AAW78189

ID AAW78189 standard; protein; 967 AA.

XX AAW78189;

AC AAW78189;

DT 13-APR-1999 (first entry)

DE Human secreted protein encoded by gene 64 clone HOUQC17.

XX Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal;
KW immune system; aschma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclasts; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

XX Homo sapiens.
OS Key Location/Qualifiers
XX Misc-difference 40 /label= unknown
FT Misc-difference 45 /label= unknown
FT Misc-difference 169 /label= unknown
FT Misc-difference 293 /label= unknown
FT Misc-difference 297 /label= unknown
FT Misc-difference 557 /label= unknown
XX WO9856804-A1.
XX 17-DEC-1998.
XX 11-JUN-1998; 98WO-US012125.
XX 13-JUN-1997; 97US-0049547P.
PR 13-JUN-1997; 97US-0049548P.
PR 13-JUN-1997; 97US-0049549P.
PR 13-JUN-1997; 97US-0049550P.
PR 13-JUN-1997; 97US-0049566P.
PR 13-JUN-1997; 97US-0049606P.
PR 13-JUN-1997; 97US-0049607P.
PR 13-JUN-1997; 97US-0049608P.
PR 13-JUN-1997; 97US-0049609P.
PR 13-JUN-1997; 97US-0049610P.
PR 13-JUN-1997; 97US-0049611P.
PR 13-JUN-1997; 97US-0050901P.
PR 13-JUN-1997; 97US-0052989P.
PR 08-JUL-1997; 97US-0051919P.
PR 18-AUG-1997; 97US-0055984P.
PR 12-SEP-1997; 97US-0058665P.
PR 12-SEP-1997; 97US-0058668P.
PR 12-SEP-1997; 97US-0058669P.
PR 12-SEP-1997; 97US-0058750P.
PR 12-SEP-1997; 97US-0058971P.
PR 12-SEP-1997; 97US-0058972P.
PR 12-SEP-1997; 97US-0058975P.
PR 02-OCT-1997; 97US-0060834P.
PR 02-OCT-1997; 97US-0060841P.
PR 02-OCT-1997; 97US-0060844P.
PR 02-OCT-1997; 97US-0060865P.
PR 02-OCT-1997; 97US-0061059P.
PR 02-OCT-1997; 97US-0061060P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Moore PA, Shi Y, Rosen CA, Ruben SM, Lafleur DW, Olsen HS;
PI Ebner R, Brewer LA, Young P, Greene JM, Ferrie AM, Yu G, Ni J;
PI Feng P;
XX MPI; 1999-080881/07.
XX N-PSDB; AAX04374.
XX New isolated human genes and the secreted polypeptides they encode -
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders.
XX Claim 11; Page 297-300; 380pp; English.
XX This sequence represents a secreted human protein encoded by the gene
CC clone detailed in the descriptor line. The gene can be used to generate
CC fusion proteins by linking to the gene to a human immunoglobulin FC
CC portion (e.g. AAX04302) for increasing the stability of the fused protein
CC as compared to the human protein only. The invention relates to 86 novel
CC genes and their fragments (nucleic acid sequences: AAX04311-X04410; amino

CC acid sequences AAW78126-W78225) which are useful for preventing, treating
 CC or ameliorating medical conditions e.g. by protein or gene therapy. Also,
 CC pathological conditions can be diagnosed by determining the amount of the
 CC new polypeptides in a sample or by determining the presence of mutations
 CC in the new polynucleotides. Specific uses are described for each of the
 CC 86 polynucleotides, based on which tissues they are most highly expressed
 CC in (see AAX04311 for described uses)
 XX
 SQ Sequence 967 AA;

Query Match 98.4%; Score 5201; DB 2; Length 967;
 Best Local Similarity 98.3%; Pred. No. 0;
 Matches 951; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY	1	MQRAVPEGFGRKLGSDMGNARAPGSRSGFVPFTLLLLAALLAVSDALGRPSDEEEL	60
DB	1	MQRAVPEGFGRKLGSDMGNARAPGSRSGFVPFTLLLLAALLAVSDALGRPSDEEEL	60
QY	61	VPELERAPGHGTTLRHLHAFDQDLDPDSSFLAPGFTLQNVGRKSGSDTLPETDL	120
DB	61	VPELERAPGHGTTLRHLHAFDQDLDPDSSFLAPGFTLQNVGRKSGSETLPETDL	120
QY	121	AHCFYSGTVNGDPSAAALSCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGEKPPAP	180
DB	121	AHCFYSGTVNGDPSAAALSCEGVRGAFYLLGEAYFIQPLPAASERLXTAAPGEKPPAP	180
QY	181	LOFHLLRRNRQDVGCTCGVVDDEPRPTGKAETDEDEGTGEDEGPQWSPQDALQGVG	240
DB	181	LOFHLLRRNRQDVGCTCGVVDDEPRPTGKAETDEDEGTGEDEGPQWSPQDALQGVG	240
QY	241	QPTGTGSIKRRKRVSSHRYVETMLVADQSMAPFSGSLKHYLLTLFVVAARLYKHPISRN	300
DB	241	QPTGTGSIKRRKRVSSHRYVETMLVADQSMAPFSGSLKHYLLTLFVVAARLYKHPXIRN	300
QY	301	SVSLVVKVILVHDEQKGPEVTSNAALTLRNCWQKOHNPSPDRDAEHYDTAILFTQQD	360
DB	301	SVSLVVKVILVHDEQKGPEVTSNAALTLRNCWQKOHNPSPDRDAEHYDTAILFTQQD	360
QY	361	LCGSQTCDTLGMADVTGCDPSRCSVTEDDGLQAAFTTAHELGHVFNMPHDDAKQACSL	420
DB	361	LCGSQTCDTLGMADVTGCDPSRCSVTEDDGLQAAFTTAHELGHVFNMPHDDAKQACSL	420
QY	421	NGVNQDSHMAASLNSLDHSPWSPCSGYMTITSFLDNGHGECCLMDKPNPIQLPGLDLPGT	480
DB	421	NGVNQDSHMAASLNSLDHSPWSPCSAYMTITSFLDNGHGECCLMDKPNPIQLPGLDLPGT	480
QY	481	SYDANRQCFITFGEDSKHCPDAASTCTLWCTGTSGGVLCVQTKHFPWADGTSCEGKWC	540
DB	481	SYDANRQCFITFGEDSKHCPDAASTCTLWCTGTSGGVLCVQTKHFPWADGTSCEGKWC	540
QY	541	INGKCVNKNHRKHEDTTPHGSWGMWGPWGDCSRCTCGGVQVYTMRECDNPVPKNGKCEG	600
DB	541	INGKCVNKTDRKHEDTTPHGSWGMWGPWGDCSRCTCGGVQVYTMRECDNPVPKNGKCEG	600
QY	601	KRVYRSCNLEDCPDNNGKTFREBQCEAHNEFSAKSGSGPAVEWIPKYAGVSPKDRCKL	660
DB	601	KRVYRSCNLEDCPDNNGKTFREBQCEAHNEFSAKSGSGPAVEWIPKYAGVSPKDRCKL	660
QY	661	ICQAKGIGYFVLQPKVVDGTPCSPDSTSVCVQGCVKAGCDRIIDSKKPKDCGVCQGN	720
DB	661	ICQAKGIGYFVLQPKVVDGTPCSPDSTSVCVQGCVKAGCDRIIDSKKPKDCGVCQGN	720
QY	721	GSTCKKISGTVSAKPGVHDIIITPTGATNLEVKORNGSRNNGSFLAIKAAAGTVILN	780
DB	721	GSTCKKISGTVSAKPGVHDIIITPTGATNLEVKORNGSRNNGSFLAIKAAAGTVILN	780
QY	781	GDYTLSTLEQIDMYKGVVLYRSGSSAALERIRSFPLKEPLTIQVLTGVGNALRPKIYTY	840
DB	781	GDYTLSTLEQIDMYKGVVLYRSGSSAALERIRSFPLKEPLTIQVLTGVGNALRPKIYTY	840
QY	841	FVKKKKGSFNAIPTFSAWVIBEWGECSSKSLGHQRLVECRDINGQPASECAKEVPAS	900
DB	841	FVKKKKGSFNAIPTFSAWVIBEWGECSSKSLGHQRLVECRDINGQPASECAKEVPAS	900

QY	901	TRPCADHPQWQWQWSSCSKTGKGYKKTSLKLSHDSGLSHDSCDPLKPKPHFIDF	960
DB	901	TRPCADHPQWQWQWSSCSKTGKGYKKTSLKLSHDSGLSHDSCDPLKPKPHFIDF	960
QY	961	CTMAECS 967	
DB	961	CTMAECS 967	

RESULT 8
 ADA57139
 ID ADA57139 standard; protein; 967 AA.
 XX
 AC ADA57139;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Human secreted protein #422.
 KW immunosuppressive; antiinflammatory; antiasthmatic; antiallergic;
 KW cyostatic; cerebroprotective; neuroprotective; nootropic;
 KW cardiovascular; antiarteriosclerotic; gene therapy;
 KW human secreted protein; immune disorder; inflammation;
 KW respiratory disorder; cancer; CNS disorder; neurodegenerative disorders;
 KW inflammatory bowel disease; nephritis; Crohn's disease; asthma; allergy;
 KW multiple sclerosis; ischaemic brain injury; Parkinson's disease;
 KW Alzheimer's disease; atherosclerosis; myocarditis; chromosome mapping;
 KW triple helix formation; antisense gene therapy; forensic biology.
 XX
 OS Homo sapiens.
 XX
 PN WO2002102994-A2.
 XX
 PD 27-DEC-2002.
 XX
 PF 19-MAR-2002; 2002WO-US008278.
 XX
 PR 21-MAR-2001; 2001US-0277340P.
 PR 19-JUL-2001; 2001US-0306171P.
 PR 13-NOV-2001; 2001US-0331287P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI; 2003-167512/16.
 DR N-PSDB; ADA56243.
 XX
 PT New human secreted polypeptides and polynucleotides, useful for
 PT diagnosing, treating or preventing e.g. immune disorders, inflammatory
 PT conditions, respiratory disorders, cancers, CNS disorders, or
 PT neurodegenerative disorders.

Claim 13; SEQ ID NO 1329; 1754pp; English.

The invention relates to 592 new human secreted polypeptides useful for
 diagnosing, treating or preventing e.g. immune disorders, inflammatory
 conditions, respiratory disorders, cancers, CNS disorders, or
 neurodegenerative disorders, or polypeptides comprising an amino acid
 sequence at least 95% identical to the new sequences. The polypeptides,
 antibodies or antibody fragments that bind to the polypeptides, nucleic
 acids encoding the polypeptides, agonists or antagonists that binds to
 the polypeptide, are useful in preparing diagnostic or pharmaceutical
 compositions for diagnosing, treating or preventing an e.g. immune
 disorders, inflammatory conditions (e.g. inflammatory bowel disease,
 nephritis or Crohn's disease), respiratory disorders (e.g. asthma and
 allergy), cancers (e.g. gastric, ovarian or lung cancer), CNS disorders
 (e.g. multiple sclerosis or ischaemic brain injury), neurodegenerative
 disorders (e.g. Parkinson's disease or Alzheimer's disease), and
 cardiovascular disorders (e.g. atherosclerosis or myocarditis). The
 polynucleotides are useful for chromosome identification, chromosome
 mapping, for controlling gene expression through triple helix formation

or antisense DNA or RNA, in gene therapy, for identifying individuals from minute biological samples, in forensic biology, and as hybridization probes. The polypeptides are useful for as molecular weight markers on sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gels, to raise antibodies, for testing biological activities, and for treating or preventing neural disorders, immune system disorders, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal, proliferative and/or cancerous diseases. This sequence corresponds to one of the polypeptide of the invention. Note: The sequence data for this patent did form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 967 AA;

Query Match 98.4%; Score 5201; DB 6; Length 967;
Best Local Similarity 98.3%; Pred. No. 0;
Matches 951; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 1 MQRVPEGFGRKLGSDMGNAERAPGSRSGFVPVTLALLAALLAVSDALGRPSEDEEL 60
DB 1 MQRVPEGFGRKLGSDMGNAERAPGSRSGFVPVTLALLAALLAVSDALGRPSEDEEL 60
QY 61 VPELERVPGHGTTLRLHAFQQLDLVPPSSFLAPGFTLQNVGRKSGSTPLPETDL 120
DB 61 VPELERVPGHGTTLRLHAFQQLDLVPPSSFLAPGFTLQNVGRKSGSTPLPETDL 120
QY 121 AHCFYSGTVNGDPSSAAALSCEGVGAPFYLGEAYFTQPLPAASERLATAAGKPPAP 180
DB 121 AHCFYSGTVNGDPSSAAALSCEGVGAPFYLGEAYFTQPLPAASERLATAAGKPPAP 180
QY 181 LQFHLRRNRQDVGTCGVVDDEPRPTGKATEDEDEGEDEGEQWSPQDPALQGVG 240
DB 181 LQFHLRRNRQDVGTCGVVDDEPRPTGKATEDEDEGEDEGEQWSPQDPALQGVG 240
QY 241 QPTGTGSIKKRFVSSHRYVETMLVADQSMABFHGSLGHYLLTLFSVAARLYKHPSTRN 300
DB 241 QPTGTGSIKKRFVSSHRYVETMLVADQSMABFHGSLGHYLLTLFSVAARLYKHPSTRN 300
QY 301 SVSLVVKLVTHDEQKPEVTSNAALTLRNCNKKOHNPPSDRDAEHYDTAILFTROD 360
DB 301 SVSLVVKLVTHDEQKPEVTSNAALTLRNCNKKOHNPPSDRDAEHYDTAILFTROD 360
QY 361 LCGSQTCTLGNADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDPAKQACSL 420
DB 361 LCGSQTCTLGNADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDPAKQACSL 420
QY 421 NGVNQDSHMMASMLNLDHSPWSPCSGYMITSFLDNGHGECLMDKPNQPIQLPGDLPGT 480
DB 421 NGVNQDSHMMASMLNLDHSPWSPCSGYMITSFLDNGHGECLMDKPNQPIQLPGDLPGT 480
QY 481 SYDANRQCFPTGERSKPCDPAASTCSTLWCTGTSGGVLVCTQKHPHADGTSCEGKWC 540
DB 481 SYDANRQCFPTGERSKPCDPAASTCSTLWCTGTSGGVLVCTQKHPHADGTSCEGKWC 540
QY 541 INGKCNKHKRKHFTDTPPHSGWMGMPGDCSRTCGGVQVYTMRECDNPVPKNGKCYCEG 600
DB 541 INGKCNKHKRKHFTDTPPHSGWMGMPGDCSRTCGGVQVYTMRECDNPVPKNGKCYCEG 600
QY 601 KRVRYSNCLDCPDNNGKTFREEQCEAHNEFSKASFGSGPAVEIPKYAGVSPKDRCKL 660
DB 601 KRVRYSNCLDCPDNNGKTFREEQCEAHNEFSKASFGSGPAVEIPKYAGVSPKDRCKL 660
QY 661 ICQAKGIGYFFVLQPKVVDGTCSPDSTSVQVQGVKAGCDRIIDSKKPKKCGVCGN 720
DB 661 ICQAKGIGYFFVLQPKVVDGTCSPDSTSVQVQGVKAGCDRIIDSKKPKKCGVCGN 720
QY 721 GSTCKKISGVTSAKPGYHDIITPTGATNIEVKORNGSRNNGSFLAIKAADGTIYN 780
DB 721 GSTCKKISGVTSAKPGYHDIITPTGATNIEVKORNGSRNNGSFLAIKAADGTIYN 780
QY 781 GDYTLSTLEQDIMYKGVVLRYSGSSAALERTSRFSPLKEPLTIQVLTGVNLRPKIKYTY 840
DB 781 GDYTLSTLEQDIMYKGVVLRYSGSSAALERTSRFSPLKEPLTIQVLTGVNLRPKIKYTY 840

Db 781 GDYTLSTLEQDIMYKGVVLRYSGSSAALERTSRFSPLKEPLTIQVLTGVNLRPKIKYTY 840
QY 841 FVKKKESFNAIPTFSAWVIEEWGECSKSCBLGQRRRLVECRDINGQPASECAKEVKPAS 900
DB 841 FVKKKESFNAIPTFSAWVIEEWGECSKSCBLGQRRRLVECRDINGQPASECAKEVKPAS 900
QY 901 TRPCADHPCPQWQJGEWSSCKTCGKGYKKTSLKCLSHDGGVLSDHSCDPLKKPKHFIDF 960
DB 901 TRPCADHPCPQWQJGEWSSCKTCGKGYKKTSLKCLSHDGGVLSDHSCDPLKKPKHFIDF 960
QY 961 CTMAECS 967
DB 961 CTMAECS 967
RESULT 9
ADA41003
ID ADA41003 standard; protein; 967 AA.
XX
AC ADA41003;
DT 20-NOV-2003 (first entry)
XX Human secreted protein.
DE Human; secreted protein; cancer; hyperproliferative disorder;
XX rheumatoid arthritis; autoimmune disorder; haematopoietic disorder;
KW anaemia; allergic reaction; asthma; cardiovascular disorder;
KW wound healing; cytostatic; immunosuppressive; neutrotropic; neuroprotective;
KW antiviral; antiallergic; hepatotropic; antidiabetic; antiinflammatory;
KW vulnery; cardiant; gene therapy.
XX Homo sapiens.
OS
XX WO2002102993-A2.
PN
XX 27-DEC-2002.
PD
XX 19-MAR-2002; 2002WO-US008123.
PF
XX 21-MAR-2001; 2001US-0277340P.
PR 19-JUL-2001; 2001US-0306171P.
PR 13-NOV-2001; 2001US-0331287P.
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Rosen CA, Ruben SM;
PI
XX WPI; 2003-175238/17.
DR
XX New human secreted proteins and nucleic acid molecules, useful for preparing a diagnostic or pharmaceutical composition for diagnosing, preventing or treating cancer or other hyperproliferative disorder, asthma, allergies or AIDS.
PT
XX Claim 1; SEQ ID NO 1385; 3205pp; English.
PS
XX The invention relates to novel genes ADA39629-ADA40565 and proteins ADA40566-ADA41501 for human secreted proteins, useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The polypeptides, nucleic acid molecules, antibodies or their fragments, and agonists or antagonists that bind to the polypeptide are useful for preparing a diagnostic or pharmaceutical composition for diagnosing or treating cancer or other hyperproliferative disorder. The polypeptides and nucleic acid molecules are also useful for detecting, preventing, diagnosing, prognosticating, treating or ameliorating cancer or other hyperproliferative disorders including neoplasms, autoimmune disorders (e.g. diabetes, rheumatoid arthritis, systemic lupus erythematosus, multiple sclerosis, autoimmune thyroiditis or haemolytic anaemia), haematopoietic or haematological disorders (e.g. anaemia, thrombocytopenia), allergic reactions including asthma or eczema, inflammatory disorders (e.g. ischaemia-reperfusion injury, inflammatory bowel disease or Crohn's disease), neurodegenerative disorders (e.g.

CC Alzheimer's disease or Parkinson's disease), cardiovascular disorders
 CC (e.g. atherosclerosis, myocarditis), infectious diseases (bacterial,
 CC fungal or viral infections including HIV/AIDS), or wound healing and
 CC disorders of epithelial cell proliferation. The nucleic acids are also
 CC useful for chromosome identification, radiation hybrid mapping or long-
 CC range restriction mapping, as molecular weight markers, or as
 CC hybridization or diagnostic probes. The polypeptides and antibodies are
 CC useful for providing immunological probes for differential identification
 CC of the tissues immunohistochemistry assays. Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX

SQ Sequence 967 AA;

Query Match 98.4%; Score 5201; DB 6; Length 967;
 Best Local Similarity 98.3%; Pred. No. 0;
 Matches 951; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 1 MORAVPEGFRKLGSDMGNAERAPGSRSGFVPTLLALLAALAVSDALGRPSDEEL 60
 DB 1 MORAVPEGFRKLGSDMGNAERAPGSRSGFVPTLLALLAALAVSDALGRPSDEEL 60
 QY 61 VPELERVPGHGTTLRLHLHAFDQDLDPDSSFLAPGFTLQNVGRKSGSDTLPETDL 120
 DB 61 VPELERAPGHGTTTLRLHLHAFDQDLDPDSSFLAPGFTLQNVGRKSGSETPLPETDL 120
 QY 121 AHCFYSGTVNGDPSSAALSCEGVGAFYLLGEAYFIQPLPAASERLATAPGEKPPAP 180
 DB 121 AHCFYSGTVNGDPSSAALSCEGVGAFYLLGEAYFIQPLPAASERLXTAPGEKPPAP 180
 QY 181 LOFHLLRRNRQDVGCTCGVVDDEPRPTGKAETDEDEGTEGEGPQSDPALQGVG 240
 DB 181 LOFHLLRRNRQDVGCTCGVVDDEPRPTGKAETDEDEGTEGEGPQSDPALQGVG 240
 QY 241 OPTGTGSRIRKRFVSSHRYVETMLVADQSMAEFFHSGGLKHYLLTLFVSAARLYKHPSTRN 300
 DB 241 OPTGTGSRIRKRFVSSHRYVETMLVADQSMAEFFHSGGLKHYLLTLFVSAARLYKHPSTRN 300
 QY 301 SVSLVVKILVHDEKQGPVTSNNAALTLRNFQWQKQHNPPSDRDAHYDTAILFTROD 360
 DB 301 SVSLVVKILVHDEKQGPVTSNNAALTLRNFQWQKQHNPPSDRDAHYDTAILFTROD 360
 QY 361 LCGSQTCDTLGADVTGCDPSRSCSVLEDGLQAAFTTAHELGHVFNMPHDDAKQACSL 420
 DB 361 LCGSQTCDTLGADVTGCDPSRSCSVLEDGLQAAFTTAHELGHVFNMPHDDAKQACSL 420
 QY 421 NGVNQDSHMASMLSNLDHSPWSPSCSAYMITSLDNGHGECMLDKPQNPIQLPGDLPGT 480
 DB 421 NGVNQDSHMASMLSNLDHSPWSPSCSAYMITSLDNGHGECMLDKPQNPIQLPGDLPGT 480
 QY 481 SYDANRQOFTFGEDSKICPDAASTCSTLWCTGTSGGVLVQCTKHPFWDATSGCGKWC 540
 DB 481 SYDANRQOFTFGEDSKICPDAASTCSTLWCTGTSGGVLVQCTKHPFWDATSGCGKWC 540
 QY 541 INGCNVNKHKHPDTPPHGSGWGMGPGDCSRCTCGGVQVYTMRECONPVKNGKCEG 600
 DB 541 INGCNVXKTDKHPDTPPHGSGWGMGPGDCSRCTCGGVQVYTMRECONPVKNGKCEG 600
 QY 601 KRVRYSCLNEDCDPNNKGTFRQCEAHNEFSAKSGSGPAVEWIPKYAGVSPKDRCKL 660
 DB 601 KRVRYSCLNEDCDPNNKGTFRQCEAHNEFSAKSGSGPAVEWIPKYAGVSPKDRCKL 660
 QY 661 ICQAKGIGYFFVLQKQVVDGTPCSPDSTSVQCGQCVKAGCDRIIDSKKFDKGVCGGN 720
 DB 661 ICQAKGIGYFFVLQKQVVDGTPCSPDSTSVQCGQCVKAGCDRIIDSKKFDKGVCGGN 720
 QY 721 GSTCKISGVSVTSAKPGVHDIITPTGATNLEVKORNGSRNNGSFLAIKAAAGTILN 780
 DB 721 GSTCKISGVSVTSAKPGVHDIITPTGATNLEVKORNGSRNNGSFLAIKAAAGTILN 780
 QY 781 GDYTLSTLEQDIMYKGVVLRYSGSSAALERIRSFSPLEKPTIQLVLTGVALRPKIKYTY 840
 DB 781 GDYTLSTLEQDIMYKGVVLRYSGSSAALERIRSFSPLEKPTIQLVLTGVALRPKIKYTY 840

DB 781 GDYTLSTLEQDIMYKGVVLRYSGSSAALERIRSFSPLEKPTIQLVLTGVALRPKIKYTY 840
 QY 841 FVKKKESFNAIPTFSAWVIEEWGECSSKSELGHQORLVECRDINGQPASECAKEVPAS 900
 DB 841 FVKKKESFNAIPTFSAWVIEEWGECSSKSELGHQORLVECRDINGQPASECAKEVPAS 900
 QY 901 TRPCADHPCPQWOLGEWSSCKTCGKYKKTSLKLSHGDGVLSHSDCDPLKPKHFIDF 960
 DB 901 TRPCADHPCPQWOLGEWSSCKTCGKYKKTSLKLSHGDGVLSHSDCDPLKPKHFIDF 960
 QY 961 CTMAECS 967
 DB 961 CTMAECS 967

RESULT 10
 ADB91631
 ID ADB91631 standard; protein; 967 AA.

XX AC ADB91631;
 XX AC ADB91631;
 DT 04-DEC-2003 (first entry)
 XX Human secreted protein #SEQ ID 577.
 DE Secreted protein; gene therapy; antidiabetic; diabetes; human.
 XX Homo sapiens.
 XX WO2003004622-A2.
 XX 16-JAN-2003.
 XX 19-MAR-2002; 2002WO-US008124.
 XX 21-MAR-2001; 2001US-0277340P.
 PR 19-JUL-2001; 2001US-0306171P.
 PR 13-NOV-2001; 2001US-0331287P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Ruben SM;
 XX WPI; 2003-229407/22.
 XX Nucleic acid encoding a human secreted protein is useful in diagnosing or
 XX treating diabetes or conditions related to diabetes.
 PS Claim 3; SEQ ID NO 577; 1537pp; English.

XX The invention relates to isolated nucleic acid molecules ADB91065-
 CC ADB91448 and ADB91835-ADB91911 encoding human secreted proteins ADB91449-
 CC ADB91834. Also disclosed is a recombinant vector comprising a
 CC polynucleotide of the invention, and a recombinant host cell comprising
 CC the recombinant vector. The polypeptide of the invention is useful in
 CC identifying a binding partner by contacting the polypeptide with a
 CC binding partner, and determining whether the binding partner increases or
 CC decreases activity of the polypeptide. The polypeptide, polynucleotide,
 CC antibody or its fragment, agonist or antagonist are useful for preparing
 CC a pharmaceutical composition for diagnosing or treating diabetes or
 CC conditions related to diabetes. The present sequence is that of the human
 CC immunoglobulin Fc portion used to generate fusion proteins, increasing
 CC the stability of the fused protein as compared to the secreted protein
 CC only. Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX

SQ Sequence 967 AA;

Query Match 98.4%; Score 5201; DB 7; Length 967;
 Best Local Similarity 98.3%; Pred. No. 0;
 Matches 951; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

```
QY 1 MORAVPEGFGRRKLGSDMGNAERAPGSRSGFVPVPTLLLLAALLAVSDALGRPSEDEEL 60
Db 1 MORAVPEGFGRRKLGSDMGNAERAPGSRSGFVPVPTLLLLAALLAVSDALGRPSEDEEL 60
QY 61 VPELERVPGHGTTRRLRHAFDQQLDLPDSSFLAPGFTLQNVGRKSGSDTLPETDL 120
Db 61 VPELERVPGHGTTRRLRHAFDQQLDLPDSSFLAPGFTLQNVGRKSGSETLPETDL 120
QY 121 AHCFYSGTVNGDPSSAAALSCEGVRGAFYLLIGEAYFTQPLPAASERLATAAPGKPPAP 180
Db 121 AHCFYSGTVNGDPSSAAALSCEGVRGAFYLLIGEAYFTQPLPAASERLATAAPGKPPAP 180
QY 181 LQFHLLRRNRQDVGCTCGVDDDEPRPTGKABTEDEDESGTEDEGPGQWSPDPAALQGVG 240
Db 181 LQFHLLRRNRQDVGCTCGVDDDEPRPTGKABTEDEDESGTEDEGPGQWSPDPAALQGVG 240
QY 241 OPTGTGSIKRRKRFVSSHRYVETMLVADQSMAEFHGSLKHYLLTLFSVAARLYKHPSIRN 300
Db 241 OPTGTGSIKRRKRFVSSHRYVETMLVADQSMAEFHGSLKHYLLTLFSVAARLYKHPSIRN 300
QY 301 SVSLVVKILVIHDEQKGPVTSNAALTILRNFQKQHNPPSDRDAEHYDTAILFTTRQD 360
Db 301 SVSLVVKILVIHDEQKGPVTSNAALTILRNFQKQHNPPSDRDAEHYDTAILFTTRQD 360
QY 361 LCGSQTCDTLGMADVGTVCDSRSRCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQKASL 420
Db 361 LCGSQTCDTLGMADVGTVCDSRSRCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQKASL 420
QY 421 NGVNQDSHMMASMLNLDHSPQWSPSCSYMYITSFLDNGHGCECLMDKPNPTQLPGDLPGT 480
Db 421 NGVNQDSHMMASMLNLDHSPQWSPSCSYMYITSFLDNGHGCECLMDKPNPTQLPGDLPGT 480
QY 481 SYDANRQCOFTFGEDSKHCPDAASTCTLWCTGTSGGVLVCOVKHPPWADGTSCEGKWC 540
Db 481 SYDANRQCOFTFGEDSKHCPDAASTCTLWCTGTSGGVLVCOVKHPPWADGTSCEGKWC 540
QY 541 INGKCVNKHRRKHFTPTPHGSGMGMWPGDCSRTCGGVQVYTMRECDNPVPKNGKCYCEG 600
Db 541 INGKCVXKTDKHFDTPTPHGSGMGMWPGDCSRTCGGVQVYTMRECDNPVPKNGKCYCEG 600
QY 601 KRVYRSCNLEDCPDNNGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKL 660
Db 601 KRVYRSCNLEDCPDNNGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKL 660
QY 661 IQAAGIGYFFVLPQKAVDGTSPDSTSVQGVQCVKAGCDRIIDSKKFKDKCGVCGGN 720
Db 661 IQAAGIGYFFVLPQKAVDGTSPDSTSVQGVQCVKAGCDRIIDSKKFKDKCGVCGGN 720
QY 721 GSTCKKISGSVTSAPGYHDIITPTGATNIEVKQNRQSRNNGSFLAIKAADGTIYLN 780
Db 721 GSTCKKISGSVTSAPGYHDIITPTGATNIEVKQNRQSRNNGSFLAIKAADGTIYLN 780
QY 781 GDYTLSTLEQDITMYKGVVLRYSGSAALERISFSPLKEPLTIQVLTGNALRPKIKTY 840
Db 781 GDYTLSTLEQDITMYKGVVLRYSGSAALERISFSPLKEPLTIQVLTGNALRPKIKTY 840
QY 841 FVKKKESFNALPTTSAWIBEWGSCSKELGWORRLVECRDINGQPASECACKVKKAS 900
Db 841 FVKKKESFNALPTTSAWIBEWGSCSKELGWORRLVECRDINGQPASECACKVKKAS 900
QY 901 TRPCADHPQWOLGEWSCKTCGKYKTKSLKCLSHDGGVLSHSDCDPLKKPKHFD 960
Db 901 TRPCADHPQWOLGEWSCKTCGKYKTKSLKCLSHDGGVLSHSDCDPLKKPKHFD 960
QY 961 CTMAECS 967
Db 961 CTMAECS 967
```

RESULT 11

ADC74267

ID ADC74267 standard; protein; 967 AA.

XX

```
AC ADC74267;
XX 01-JAN-2004 (first entry)
XX Human secreted protein - SEQ ID 900.
XX antianemic; antirheumatic; antiarthritic; antiinflammatory; antithyroid;
XX antidabetic; immunosuppressive; dermatologic; nephrotropic;
XX antiparkinsonian; neuroprotective; nootropic; antibacterial; virucide;
XX fungicide; antiparasitic; antiarteriosclerotic; vulnerary; cytostatic;
XX haemopoietic; haematologic; anaemia; autoimmune disorder;
XX rheumatoid arthritis; inflammation; Grave's disease; diabetes;
XX systemic lupus erythematosus; glomerulonephritis; neurodegenerative;
XX Parkinson's; Alzheimer's; wound; hyperproliferative; atherosclerosis;
XX cancer; bacterial; viral; fungal; parasitic infection; gene therapy;
XX human.
XX Homo sapiens.
OS WO2003038063-A2.
XX 08-MAY-2003.
XX 19-MAR-2002; 2002WO-US008277.
XX 21-MAR-2001; 2001US-0277340P.
XX 19-JUL-2001; 2001US-0306171P.
XX 13-NOV-2001; 2001US-0331287P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Ruben SM;
XX WPI: 2003-430516/40.
XX N-PSDB; ADC73652.
XX New human secreted polypeptide for diagnosing, preventing or treating
XX hemopoietic or hematologic disorders (e.g. anemia), autoimmune
XX disorders (e.g. diabetes) or hyperproliferative disorders (e.g. cancer or
XX atherosclerosis).
XX Claim 16; SEQ ID NO 900; 2272pp; English.
XX The invention relates to a novel human secreted polypeptide comprising a
XX defined sequence given in the specification. The polypeptide, nucleic
XX acid molecule, antibody, agonist or antagonist of the invention may be
XX useful for preparing a composition for diagnosing or treating a
XX haemopoietic or haematologic disorder such as anaemia, autoimmune
XX disorders such as rheumatoid arthritis, inflammation, Grave's disease,
XX diabetes, systemic lupus erythematosus or glomerulonephritis,
XX neurodegenerative disorders including Parkinson's disease and Alzheimer's
XX disease, wounds and hyperproliferative disorders including
XX atherosclerosis or cancer, as well as bacterial, viral, fungal or
XX parasitic infections. The polypeptide may also be used during gene
XX therapy procedures and for identifying a binding partner by contacting
XX the polypeptide with a binding partner and determining whether the
XX binding partner increases or decreases the activity of the polypeptide.
XX The current sequence is that of the human secreted protein of the
XX invention.
XX Sequence 967 AA;
```

Query Match 98.4%; Score 5201; DB 7; Length 967;

Best Local Similarity 98.3%; Pred. No. 0;

Matches 951; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 1 MORAVPEGFGRRKLGSDMGNAERAPGSRSGFVPVPTLLLLAALLAVSDALGRPSEDEEL 60

Db 1 MORAVPEGFGRRKLGSDMGNAERAPGSRSGFVPVPTLLLLAALLAVSDALGRPSEDEEL 60

QY 61 VPELERVPGHGTTRRLRHAFDQQLDLPDSSFLAPGFTLQNVGRKSGSDTLPETDL 120

Db 61 VPELERVPGHGTTRRLRHAFDQQLDLPDSSFLAPGFTLQNVGRKSGSETLPETDL 120

Qy	121	AHCFYSGTVNGDPSSAALSICEGYRGAFYLLGEAYFIQPLPAASERLATAAPEKPPAP	180
Db	121	AHCFYSGTVNGDPSSAALSICEGYRGAFYLLGEAYFIQPLPAASERLATAAPEKPPAP	180
Qy	181	LQFHLLRRNRQDVGCTCGVVDDEPRPTGKAETEDDEGTEDEGEGPQWSPQDPALQGVG	240
Db	181	LQFHLLRRNRQDVGCTCGVVDDEPRPTGKAETEDDEGTEDEGEGPQWSPQDPALQGVG	240
Qy	241	QPTGTGSIRKKRFVSSHRYVETMLVADQSMAEFHGSLKHLYLLTLFSSVAARLYKHPSIRN	300
Db	241	QPTGTGSIRKKRFVSSHRYVETMLVADQSMAEFHGSLKHLYLLTLFSSVAARLYKHPSIRN	300
Qy	301	SVSLVWVKILVIHDEQKPEVTSNAALTLRNECNWQKOHNPSPDRDAEHYDTAILFTROD	360
Db	301	SVSLVWVKILVIHDEQKPEVTSNAALTLRNECNWQKOHNPSPDRDAEHYDTAILFTROD	360
Qy	361	LCGSQTCDTLGMADVTCVDPSSRSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQCASL	420
Db	361	LCGSQTCDTLGMADVTCVDPSSRSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQCASL	420
Qy	421	NGVNQDSHMMAISLNLNLDHSPFSPCGYMITSFILNDHGHCIMDKKPNQPIQLPGDLPGT	480
Db	421	NGVNQDSHMMAISLNLNLDHSPFSPCGYMITSFILNDHGHCIMDKKPNQPIQLPGDLPGT	480
Qy	481	SYDANRQCOFTFGEDSKHCPDAASTCTLWCTGTSGGVLVCOTKHPFWADGTSCEGKWC	540
Db	481	SYDANRQCOFTFGEDSKHCPDAASTCTLWCTGTSGGVLVCOTKHPFWADGTSCEGKWC	540
Qy*	541	INGKCVNKNRKHFDTPFHGSGMWGMPGDCSRTCCTGGGVQVYTMRECONVPVKNKGKCEG	600
Db	541	INGKCVNKNRKHFDTPFHGSGMWGMPGDCSRTCCTGGGVQVYTMRECONVPVKNKGKCEG	600
Qy	601	KRVYRSCNLEDCPDNNGKTPREOQEAHNEFSAKSGSGPAVEWIPKYAGVSPKORCKL	660
Db	601	KRVYRSCNLEDCPDNNGKTPREOQEAHNEFSAKSGSGPAVEWIPKYAGVSPKORCKL	660
Qy	661	ICQAKGIGYFFVLQPKVVDGTPCSPDSTSVCCVQGCYKAGCDRIIDSKKFDKGCVCGN	720
Db	661	ICQAKGIGYFFVLQPKVVDGTPCSPDSTSVCCVQGCYKAGCDRIIDSKKFDKGCVCGN	720
Qy	721	GSTCKKISGVSATKPGVHDIITPTGATNIEVKQRNQRGSRNNGSFLAIKAADGTIILN	780
Db	721	GSTCKKISGVSATKPGVHDIITPTGATNIEVKQRNQRGSRNNGSFLAIKAADGTIILN	780
Qy	781	GDYTLSTLEQIMYKGVVLRYSGSSAALERIRSPSPLKEPLTIQVLTGVGNALRPKIYTY	840
Db	781	GDYTLSTLEQIMYKGVVLRYSGSSAALERIRSPSPLKEPLTIQVLTGVGNALRPKIYTY	840
Qy	841	FVKKKESFNAPTFSAWVIEWGECSKSCELGWORLVECRDINGOPASECAKEVPKAS	900
Db	841	FVKKKESFNAPTFSAWVIEWGECSKSCELGWORLVECRDINGOPASECAKEVPKAS	900
Qy	901	TRPCADHPCPQWQJGEWSSCKTCGKYKKTSLKCLSHDGGVLGSHDSCDPLKPKHFIDF	960
Db	901	TRPCADHPCPQWQJGEWSSCKTCGKYKKTSLKCLSHDGGVLGSHDSCDPLKPKHFIDF	960
Qy	961	CTMAECS 967	
Db	961	CTMAECS 967	
RESULT 12			
ADD37948			
ID	ADD37948	standard; protein; 967 AA.	
XX	AC	ADD37948;	
XX	XX		
DT	15-JAN-2004	(first entry)	
XX	XX		
DE	Human	secreted protein #131.	
XX	XX	human	secreted protein; Antiallergic; Antiinflammatory; Antibacterial;
KW			

KW	Anti-HIV; Cytostatic; Immunosuppressive; Hemostatic.		
XX			
OS	Homo sapiens.		
XX			
FN	WO200290526-A2.		
XX			
PD	14-NOV-2002.		
XX			
PF	19-MAR-2002; 2002WO-US008279.		
XX			
PR	21-MAR-2001; 2001US-0277340P.		
PR	19-JUL-2001; 2001US-0306171P.		
PR	13-NOV-2001; 2001US-0331287P.		
XX			
FA	(HUMA-) HUMAN GENOME SCI INC.		
XX			
PI	Rosen CA, Ruben SM;		
XX			
DR	WPI; 2003-140218/13..		
XX			
PT	New human secreted proteins and nucleic acid molecules, useful for preparing a diagnostic or pharmaceutical composition for diagnosing or treating allergic or asthmatic disorders, or related immediate hypersensitivity disorders.		
PT			
PT			
XX			
PS	Claim 1; SEQ ID NO 430; 1323pp; English.		
XX			
CC	The present invention relates to an isolated polypeptide or human secreted protein. The polypeptides, nucleic acid molecules, antibodies or their fragments, and agonists or antagonists that bind are useful for preparing a diagnostic or pharmaceutical composition for diagnosing or treating allergic or asthmatic disorders. The polypeptide is also useful for identifying a binding partner by contacting the polypeptide with a binding partner, and determining whether the binding partner increases or decreases the activity of the polypeptide. The polypeptides and nucleic acid molecules are also useful for detecting, preventing, diagnosing, prognosticating, treating or ameliorating inflammatory disorders		
CC	neoplastic diseases , wound healing and disorders of epithelial cell proliferation, immune disorders , cardiovascular disorders, blood-related disorders, infectious diseases, endocrine disorders , or gastrointestinal disorders. The nucleic acids are also useful for chromosome identification, radiation hybrid mapping or long-range restriction mapping, as molecular weight markers, or as hybridization or diagnostic probes. The polypeptides and antibodies are useful for providing immunological probes for differential identification of the tissues		
CC	immunohistochemistry assays. The present sequence represents a human secreted protein.		
XX			
SQ	Sequence 967 AA;		
	Query Match 98.4%; Score 5201; DB 7; Length 967;		
	Best Local Similarity 98.3%; Pred. No. 0;		
	Matches 951; Conservative 4; Mismatches 12; Indels 0; Gaps 0;		
Qy	1	MQRAVPEGFRRKLGSDMGNAERAPGSRSGFVPPTLLLLAAALLAVSDALGRPSDEEL	60
Db	1	MQRAVPEGFRRKLGSDMGNAERAPGSRSGFVPPTLLLLXALLXVDALGRPSDEEL	60
Qy	61	VPELERVPGHGTRRLRLHAFDQDLDDVPDSSFLAPGFTLQNVGRKSGSDTLPETDL	120
Db	61	VPELERAPGHGTRRLRLHAFDQDLDDVPDSSFLAPGFTLQNVGRKSGSETLPETDL	120
Qy	121	AHCFYSGTVNGDPSSAAALSCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGKPPAP	180
Db	121	AHCFYSGTVNGDPSSAAALSCEGVRGAFYLLGEAYFIQPLPAASERLXTAAPGKPPAP	180
Qy	181	LQFHLLRRNRQDVGCTCGVVDDEPRPTGKAETEDDEGTEDEGPGQWSPQDALQGVG	240
Db	181	LQFHLLRRNRQDVGCTCGVVDDEPRPTGKAETEDDEGTEDEGPGQWSPQDALQGVG	240
Qy	241	QPTGTGSIRKKRFVSSHRYVETMLVADQSMAPFHSGSLKHLYLLTLFSSVAARLYKHPSIRN	300
Db	241	QPTGTGSIRKKRFVSSHRYVETMLVADQSMAPFHSGSLKHLYLLTLFSSVAARLYKHPSIRN	300

QY 301 SVSLVVVKILVTHDEQKGEVTSNAALTLRNFQCNWQKHNPSPDRDAEHYDTAILFTROD 360
DB 301 SVSLVVVKILVTHDEQKGEVTSNAALTLRNFQCNWQKHNPSPDRDAEHYDTAILFTROD 360
QY 361 LCGSOTCDTLGWADVGTVCDPSPSCSVIEDDGLQAAFTTAHELGHVFNPHDQKQASL 420
DB 361 LCGSOTCDTLGWADVGTVCDPSPSCSVIEDDGLQAAFTTAHELGHVFNPHDQKQASL 420
QY 421 NGVNQDSHMAWSMLNLDHSDQSPSCGYMITSFLDNGHGECLMDKPNPQIOLPGDLPGT 480
DB 421 NGVNQDSHMAWSMLNLDHSDQSPSCGYMITSFLDNGHGECLMDKPNPQIOLPGDLPGT 480
QY 481 SYDANRQCFTFGEDSKHCPDAASTCSTLWCTGTSGVLVVCQTKHPFWADGTSCEGKWC 540
DB 481 SYDANRQCFTFGEDSKHCPDAASTCSTLWCTGTSGVLVVCQTKHPFWADGTSCEGKWC 540
QY 541 INKCVNKHRRKHFDTPFHSGWGMWGPWGDGSRCTCGGVGYTMRCDNVPKNGKCYCG 600
DB 541 INKCVNKHRRKHFDTPFHSGWGMWGPWGDGSRCTCGGVGYTMRCDNVPKNGKCYCG 600
QY 601 KRVYRSCNLEDCPDNNGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKL 660
DB 601 KRVYRSCNLEDCPDNNGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKL 660
QY 661 ICOAKGIGYFFVLQPKVVDGTPCSPDSTSVQVQGVKAGCDRIIDSKKKFDKCGVCGGN 720
DB 661 ICOAKGIGYFFVLQPKVVDGTPCSPDSTSVQVQGVKAGCDRIIDSKKKFDKCGVCGGN 720
QY 721 GSTCKKISGVSATSAKPGYHDIITPTGATNIEVKQNRQSRNNGSFLAIAKADGTIILN 780
DB 721 GSTCKKISGVSATSAKPGYHDIITPTGATNIEVKQNRQSRNNGSFLAIAKADGTIILN 780
QY 781 GYTLSTLEQDLMYKGVLYRGSSAALRIISFSPLEPLTIQVLTGVALRPKIKYTY 840
DB 781 GYTLSTLEQDLMYKGVLYRGSSAALRIISFSPLEPLTIQVLTGVALRPKIKYTY 840
QY 841 FVKKKESFNAIPTFSAMVIEBWGECSKSELGQRRRLVECRDINGQPASECAKEYKPAS 900
DB 841 FVKKKESFNAIPTFSAMVIEBWGECSKSELGQRRRLVECRDINGQPASECAKEYKPAS 900
QY 901 TRPCADHPCPQWLGWSSCSKTCGKYKTSKLSLHSHDGGVLSHDSCDPLKKPHFIDF 960
DB 901 TRPCADHPCPQWLGWSSCSKTCGKYKTSKLSLHSHDGGVLSHDSCDPLKKPHFIDF 960
QY 961 CTMAECS 967
DB 961 CTMAECS 967
RESULT 13
AAY49501
ID AAY49501 standard; protein; 950 AA.
XX
AC AAY49501;
XX
DT 10-JAN-2000 (first entry)
XX
DE Human METH1 protein.
XX
KW Human; METH1; METH2; anti-angiogenic; metalloprotease thrombospondin;
KW cancer; diagnosis; hyperproliferative disorder; autoimmune disease;
KW angiogenesis inhibitor; abnormal wound healing; inflammation;
KW rheumatoid arthritis; psoriasis; endometrial bleeding disorder;
KW diabetic retinopathy; macula degeneration; haemangioma; detection;
KW arterial-venous malformation; immune deficiency.
XX
OS Homo sapiens.
XX
PN WO9937660-A1.
XX
PD 29-JUL-1999.
XX

PF 22-JAN-1999; 99WO-US001313.
XX
PR 23-JAN-1998; 98US-0072298P.
PR 28-AUG-1998; 98US-0098539P.
XX
PA (IRUE/) IRUELA-ARISPE L.
PA (HAST/) HASTINGS G A.
PA (RUBE/) RUBEN S M.
XX
PI Iruela-Arispe L, Hastings GA, Ruben SM;
XX
XX
DR WPI; 1999-590684/50.
DR N-PSDB; AAZ32000.
XX
PT New isolated metalloprotease thrombospondin polypeptides, useful for
PT treating hyperproliferative disorders, cancers or autoimmune disorders.
XX
PS Claim 10; Fig 1; 457pp; English.
XX
CC AAZ32000 and AAZ32001 encode, and AAY49501 and AAY49502 represent, human
CC metalloprotease thrombospondin (METH) proteins METH1 and METH2
CC respectively. METH1 and METH2 have been found to be potent inhibitors of
CC angiogenesis both in vitro and in vivo. They can be used for treating
CC cancer and other disorders related to angiogenesis including abnormal
CC wound healing, inflammation, rheumatoid arthritis, psoriasis, endometrial
CC bleeding disorders, diabetic retinopathy, some forms of macula
CC degeneration, haemangiomas, and arterial-venous malformations. They may
CC be useful in treating deficiencies or disorders of the immune system, by
CC activating or inhibiting the proliferation, differentiation, or
CC mobilisation (chemotaxis) of immune cells. The etiology of these immune
CC deficiencies or disorders may be genetic, somatic, such as cancer or some
CC autoimmune disorders, acquired (e.g. by chemotherapy or toxins), or
CC infectious. They can also be used to treat inflammatory conditions, both
CC chronic and acute conditions. The products can also be used for detection
CC and diagnosis. AAZ32002 to AAZ32080, and AAY49503 to AAY49511 represent
CC sequences given in the exemplification of the present invention
XX
SQ Sequence 950 AA;
Query Match 97.3%; Score 5144; DB 2; Length 950;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 940; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
QY 18 MGNAPRPGSRSGFPGVPTLLALLAALLAVSDALGRPSSEDEELVPELVRPGHCTTLR 77
DB 1 MGNAPRPGSRSGFPGVPTLLALLAALLAVSDALGRPSSEDEELVPELVRPGHCTTLR 60
QY 78 LHAFOQLDLVPPDSSFLAPGFTLQNVGRKSGSTPLPETDLAHCFSYGTNGDPSSAA 137
DB 61 LHAFOQLDLVPPDSSFLAPGFTLQNVGRKSGSTPLPETDLAHCFSYGTNGDPSSAA 120
QY 138 ALSLCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGKPPAPLOFHLRRNRQDVGDT 197
DB 121 ALSLCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGKPPAPLOFHLRRNRQDVGDT 180
QY 198 CGVVDDEPRPTGKAETDEDEGEDEGPQWSPQDPALQGVQPTGTGSRKRFVSSH 257
DB 181 CGVVDDEPRPTGKAETDEDEGEDEGPQWSPQDPALQGVQPTGTGSRKRFVSSH 240
QY 258 RYVETMLVADQSMABFHSGSLKHYLLTLFSVAARLYKHPSIRNSVSLVVKILVTHDEQK 317
DB 241 RYVETMLVADQSMABFHSGSLKHYLLTLFSVAARLYKHPSIRNSVSLVVKILVTHDEQK 300
QY 318 GPEVTSNAALTLRNFQCNWQKHNPSPDRDAEHYDTAILFTQDLCSGTCDTLGMADVGT 377
DB 301 GPEVTSNAALTLRNFQCNWQKHNPSPDRDAEHYDTAILFTQDLCSGTCDTLGMADVGT 360
QY 378 VCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDQKQASLNGVNDSHMAWSMLN 437
DB 361 VCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDQKQASLNGVNDSHMAWSMLN 420
QY 438 DHSQWSPSCGYMITSFLDNGHGECLMDKPNPQIOLPGDLPGTSDANRQCFTFGEDSK 497
DB

Db 421 DHSQWSPSCAYMITSFLDNGHGECLMDKPNQIQLPGDLPGTSDANRQCFTEGDSK 480
 QY 498 HCPDAASTCTLWCTGTSGGVLVCTQKHPADGTSCGEGKWCINGKCVNKHRRKHFDTP 557
 Db 481 HCPDAASTCTLWCTGTSGGVLVCTQKHPADGTSCGEGKWCINGKCVNKHRRKHFDTP 540
 QY 558 FHSGMGMWPGWDCSRTCGGGVQYTMRECDNVPKNGGKYCEGKRVYRSCNLEDCPDNN 617
 Db 541 FHSGMGMWPGWDCSRTCGGGVQYTMRECDNVPKNGGKYCEGKRVYRSCNLEDCPDNN 600
 QY 618 GKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLIQCAKGIGYFFVLQPKV 677
 Db 601 GKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLIQCAKGIGYFFVLQPKV 660
 QY 678 VDGTPCSPDSTVVCVQGCVKAGCDRIIDSKKKFKDCKGVCNGGSGTCKKISGSVTSAPG 737
 Db 661 VDGTPCSPDSTVVCVQGCVKAGCDRIIDSKKKFKDCKGVCNGGSGTCKKISGSVTSAPG 720
 QY 738 YHDIITPTGATNIEVKQNRQSRNNGSFLAIKAADGTYLNGDYTLSTLEQDLMYKGV 797
 Db 721 YHDIITPTGATNIEVKQNRQSRNNGSFLAIKAADGTYLNGDYTLSTLEQDLMYKGV 780
 QY 798 VLRYSGSSAALERIRSFSPLEPLTIQVLTVGNALRPKI KYTFYVKKKESFNAIPTFSA 857
 Db 781 VLRYSGSSAALERIRSFSPLEPLTIQVLTVGNALRPKI KYTFYVKKKESFNAIPTFSA 840
 QY 858 WVEIEMGECSKSELGWQRRLVECRDINGQPASECAKEVKPASTRPCADHPCPQWLGEW 917
 Db 841 WVEIEMGECSKSELGWQRRLVECRDINGQPASECAKEVKPASTRPCADHPCPQWLGEW 900
 QY 918 SSCSKTCGKYKTSKCLSHDGGVLSHDSCDPLKPKHFDICTMAECS 967
 Db 901 SSCSKTCGKYKTSKCLSHDGGVLSHDSCDPLKPKHFDICTMAECS 950

RESULT 14

AAAB73549
 ID AAAB73549 standard; protein; 950 AA.

XX AAAB73549;

XX 07-AUG-2001 (first entry)

XX Human ADAM-type metalloprotease MDTs4, SEQ ID NO:4.

XX Human; MDTs4; ADAM-type metalloprotease; drug screening;
 KW A Disintegrin And Metalloprotease; cancer; arthritis.

XX Homo sapiens.

XX JP2001017183-A.

XX 23-JAN-2001.

XX 09-JUL-1999; 99JP-00196584.

XX 09-JUL-1999; 99JP-00196584.

XX (YAMA) YAMANOUCHI PHARM CO LTD.

XX WPI; 2001-275950/29.

XX N-PSDB; AAH20224.

XX A new metal protease and its preparation for use as an anti-cancer and
 PT anti-arthritis therapeutic.

XX Claim 1; Page 12-14; 22pp; Japanese.

XX The invention relates to the novel human ADAM (A Disintegrin And

CC Metalloprotease)-type metalloproteases MDTs4 (AAAB73549) and MDTs5
 CC (AAAB73550). The metalloproteases can be used for the treatment of cancers
 CC and arthritis. The invention also relates to the genes encoding MDTs4 and
 CC MDTs5, vectors and host cells containing the MDTs4 or MDTs5 genes, the

CC recombinant production of MDTs4 and MDTs5, and antibody specific for
 CC MDTs4 or MDTs5, and methods of screening for compounds which modulate the
 CC activity of MDTs4 and/or MDTs5. The present sequence represents human
 CC MDTs4
 XX
 SQ Sequence 950 AA;

Query Match Similarity 97.3%; Score 5144; DB 4; Length 950;
 Best Local Similarity 98.9%; Pred. No. 0;
 Matches 940; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 18 MGNERAPGSRFGPVPTLLLLAALLAVSDALGRPSEDEELVPELERVPGHGTTLRL 77
 Db 1 MGNERAPGSRFGPVPTLLLLAALLAVSDALGRPSEDEELVPELERVPGHGTTLRL 60
 QY 78 LHAFDQQLDLVPPDSSFLAPGFTLQNVGRKSGSDTLPETDLAHCFYSGTVNGDPSAA 137
 Db 61 LHAFDQQLDLRLPDSSFLAPGFTLQNVGRKSGSETLPETDLAHCFYSGTVNGDPSAA 120
 QY 138 ALSICEGVRGAFYLLGEAYFIQPLPAASERLATAPGEKPPAPLQFHLIRNRQDYGCT 197
 Db 121 ALSICEGVRGAFYLLGEAYFIQPLPAASERLATAPGEKPPAPLQFHLIRNRQDVGCT 180
 QY 198 CGVVDDPRPTGKAETDEDEGTEGEGPQWSPQDPALQGVQPTGTGSIKKRFFVSSH 257
 Db 181 CGVVDDPRPTGKAETDEDEGTEGEGPQWSPQDPALQGVQPTGTGSIKKRFFVSSH 240
 QY 258 RYVETMLVDAQSMAEPHSGGLKHYLLTLFVSAARLYKHPISIRNSVLSLVVVKILVHDEQK 317
 Db 241 RYVETMLVDAQSMAEPHSGGLKHYLLTLFVSAARLYKHPISIRNSVLSLVVVKILVHDEQK 300
 QY 318 GPEVTSNAALTLRNFCNWKOHNPSPDRDAEHYDTAILFTRQDLCSGOTCDTLGMADVGT 377
 Db 301 GPEVTSNAALTLRNFCNWKOHNPSPDRDAEHYDTAILFTRQDLCSGOTCDTLGMADVGT 360
 QY 378 VCDPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACASLNGVQDSHMASLSNL 437
 Db 361 VCDPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACASLNGVQDSHMASLSNL 420
 QY 438 DHSQWSPSCGYMITSFLDNGHGECLMDKPNQIQLPGDLPGTSDANRQCFTEGDSK 497
 Db 421 DHSQWSPSCGYMITSFLDNGHGECLMDKPNQIQLPGDLPGTSDANRQCFTEGDSK 480
 QY 498 HCPDAASTCTLWCTGTSGGVLVCTQKHPADGTSCGEGKWCINGKCVNKHRRKHFDTP 557
 Db 481 HCPDAASTCTLWCTGTSGGVLVCTQKHPADGTSCGEGKWCINGKCVNKHRRKHFDTP 540
 QY 558 FHSGMGMWPGWDCSRTCGGGVQYTMRECDNVPKNGGKYCEGKRVYRSCNLEDCPDNN 617
 Db 541 FHSGMGMWPGWDCSRTCGGGVQYTMRECDNVPKNGGKYCEGKRVYRSCNLEDCPDNN 600
 QY 618 GKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLIQCAKGIGYFFVLQPKV 677
 Db 601 GKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLIQCAKGIGYFFVLQPKV 660
 QY 678 VDGTPCSPDSTVVCVQGCVKAGCDRIIDSKKKFKDCKGVCNGGSGTCKKISGSVTSAPG 737
 Db 661 VDGTPCSPDSTVVCVQGCVKAGCDRIIDSKKKFKDCKGVCNGGSGTCKKISGSVTSAPG 720
 QY 738 YHDIITPTGATNIEVKQNRQSRNNGSFLAIKAADGTYLNGDYTLSTLEQDLMYKGV 797
 Db 721 YHDIITPTGATNIEVKQNRQSRNNGSFLAIKAADGTYLNGDYTLSTLEQDLMYKGV 780
 QY 798 VLRYSGSSAALERIRSFSPLEPLTIQVLTVGNALRPKI KYTFYVKKKESFNAIPTFSA 857
 Db 781 VLRYSGSSAALERIRSFSPLEPLTIQVLTVGNALRPKI KYTFYVKKKESFNAIPTFSA 840
 QY 858 WVEIEMGECSKSELGWQRRLVECRDINGQPASECAKEVKPASTRPCADHPCPQWLGEW 917
 Db 841 WVEIEMGECSKSELGWQRRLVECRDINGQPASECAKEVKPASTRPCADHPCPQWLGEW 900
 QY 918 SSCSKTCGKYKTSKCLSHDGGVLSHDSCDPLKPKHFDICTMAECS 967
 Db 901 SSCSKTCGKYKTSKCLSHDGGVLSHDSCDPLKPKHFDICTMAECS 950

Db 901 SSCSKTCGKYKRSLSKLSHDSGVLSHESCDPLKKPKHPIDFCTMAECS 950

RESULT 15

AAB50002
ID AAB50002 standard; protein; 950 AA.

XX AC AAB50002;

XX DT 19-MAR-2001 (first entry)

XX DE Human METH1.

XX KW Human; METH1; metalloprotease; thrombospondin; angiogenesis inhibition;
KW cancer therapy; benign tumour; ocular angiogenic disease;
KW rheumatoid arthritis; psoriasis; wound healing; endometriosis;
KW vasculogenesis; granulation; hypertrophic scar; nonunion fracture;
KW scleroderma, trachoma; vascular adhesion; myocardial angiogenesis;
KW coronary collateral; cerebral collateral; arteriovenous malformation;
KW ischaemic limb angiogenesis; Osler-Webber syndrome;
KW plaque neovascularisation; telangiectasia; haemophilic joint;
KW angiofibroma; fibromuscular dysplasia; wound granulation;
KW Crohn's disease; atherosclerosis; birth control.

XX OS Homo sapiens.

XX PN WO200071577-A1.

XX PD 30-NOV-2000.

XX PF 25-MAY-2000; 2000WO-US014462.

XX PR 25-MAY-1999; 99US-00318208.

XX PR 20-JUL-1999; 99US-0144882P.

XX PR 10-AUG-1999; 99US-0147823P.

XX PR 13-AUG-1999; 99US-00373658.

XX PR 22-DEC-1999; 99US-0171503P.

XX PR 22-FEB-2000; 2000US-0183792P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX PA (SMITK) SMITKLINE BEECHAM CORP.

XX PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.

XX PA (IRUE/) IRUELA-ARISPE L.

XX PA (HAST-) HASTINGS G A.

XX PA (RUBE/) RUBEN S M.

XX PA (JONA/) JONAK Z L.

XX PA (TRUL/) TRULLI S H.

XX PA (FORN/) FORNWALD J A.

XX PA (TERR/) TERRETT J A.

XX PI Iruela-Arispe L, Hastings GA, Ruben SM, Jonak ZL, Trulli SH;

XX PI Fornwald JA, Terrett JA;

XX DR WPI; 2001-025136/03.

XX DR N-PSDB; AAC90057.

XX METH1 and METH2 polynucleotides and encoded polypeptides, used to inhibit
XX angiogenesis in the treatment of disorders such as cancer, rheumatoid
XX arthritis and psoriasis.

XX Claim 15; Fig 1; 768pp; English.

XX The present sequence is human METH1 (ME for metalloprotease and TH for
XX thrombospondin). METH1 can be used for inhibiting angiogenesis in an
XX individual, and for treating cancer, benign tumours, an ocular angiogenic
XX disease, rheumatoid arthritis, psoriasis, delayed wound healing,
XX endometriosis, vasculogenesis, granulations, hypertrophic scars, nonunion
XX fractures, scleroderma, trachoma, vascular adhesions, myocardial
XX angiogenesis, coronary collateral, cerebral collateral, arteriovenous
XX malformations, ischaemic limb angiogenesis, Osler-Webber syndrome, plaque
XX neovascularisation, telangiectasia, haemophilic joints, angiofibroma,
XX fibromuscular dysplasia, wound granulation, Crohn's disease or
XX atherosclerosis. METH1 can also be used in birth control. METH1 can also

CC be used in diagnostic methods for the prognosis of cancer

XX

SQ Sequence 950 AA;

Query Match 97.3%; Score 5144; DB 4; Length 950;

Best Local Similarity 98.9%; Pred. No. 0;

Matches 940; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 18 MGNAERAPGSRSGFVPPTLLALLAALAVSDALGRPSEDEBELVVPPELVRVPGHGTTLRL 77

DB 1 MGNAERAPGSRSGFVPPTLLALLAALAVSDALGRPSEDEBELVVPPELVRVPGHGTTLRL 60

QY 78 LHAFOQDLDDVPDSSFLAFQFTLQNVGRKSGSDTLPETDLAHCFSVGTVNGDPSSAA 137

DB 61 LHAFOQDLLELRDSSFLAFQFTLQNVGRKSGSETPLPETDLAHCFSVGTVNGDPSSAA 120

QY 138 ALSLCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGKPPAPLQFHLRLNRQGDVGCT 197

DB 121 ALSLCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGKPPAPLQFHLRLNRQGDVGCT 180

QY 198 CGVVDDEPRPTGKAETDEDETEGEDEGPQWSPQDPALQGVGQPTGTGSIKKRFVSSH 257

DB 181 CGVVDDEPRPTGKAETDEDETEGEDEGPQWSPQDPALQGVGQPTGTGSIKKRFVSSH 240

QY 258 RYVETMLVADQSMARFHSGLKHYLLTLFSVAARLYKHPISIRNSVSLVVKILVIHDSQK 317

DB 241 RYVETMLVADQSMARFHSGLKHYLLTLFSVAARLYKHPISIRNSVSLVVKILVIHDSQK 300

QY 318 GPEVTSNAALTLRNFCNWKQHNPPSDRDAHYDTAILFTQDLGCGSOTCDTLGNADVGT 377

DB 301 GPEVTSNAALTLRNFCNWKQHNPPSDRDAHYDTAILFTQDLGCGSOTCDTLGNADVGT 360

QY 378 VCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKQCCASLNGVNDOSHMMASLNL 437

DB 361 VCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKQCCASLNGVNDOSHMMASLNL 420

QY 438 DHSQPWSPSCGYMITSFLDNGHGECLMDKPNQPTQLPGDLPGTSDANRQCQFTFGEDSK 497

DB 421 DHSQPWSPSCGYMITSFLDNGHGECLMDKPNQPTQLPGDLPGTSDANRQCQFTFGEDSK 480

QY 498 HCPDAASTCSTLWCTGTSGGVLCQTKHFPWADGTCGEGKWCINGKCNKNHRKHFTDP 557

DB 481 HCPDAASTCSTLWCTGTSGGVLCQTKHFPWADGTCGEGKWCINGKCNKNHRKHFTDP 540

QY 558 FHSGWGMGPMGDCSRTCGGVQYTMRECDNPVPKNGGKYCEGKRVYRSCNLEDCPDNN 617

DB 541 FHSGWGMGPMGDCSRTCGGVQYTMRECDNPVPKNGGKYCEGKRVYRSCNLEDCPDNN 600

QY 618 GKTFREEQCEAHNBFSSKASFGSGPAVEWIPKYAGVSPKDRCKLICQAKGIGYFFVLQPKV 677

DB 601 GKTFREEQCEAHNBFSSKASFGSGPAVEWIPKYAGVSPKDRCKLICQAKGIGYFFVLQPKV 660

QY 678 VDGTPCSPDSTVCQGVQKAGCDRIIDSKKFKDCKGCVCGNGSTCKKISGVS TSAKPG 737

DB 661 VDGTPCSPDSTVCQGVQKAGCDRIIDSKKFKDCKGCVCGNGSTCKKISGVS TSAKPG 720

QY 738 YHDIITITPTGATNIEVKQNRQGRNNGSFLAIKAADGTIYLNGDYTLSTLEQDIMYKGV 797

DB 721 YHDIITITPTGATNIEVKQNRQGRNNGSFLAIKAADGTIYLNGDYTLSTLEQDIMYKGV 780

QY 798 VLRYSGSSAALERIRSFSPLEPLTIQVLTVGNALRPKIKTYTFVKKKESFNAIPTFSA 857

DB 781 VLRYSGSSAALERIRSFSPLEPLTIQVLTVGNALRPKIKTYTFVKKKESFNAIPTFSA 840

QY 858 WVIIEWGSCSKCELGWRRVLECRDINGQPASCAKEVKPASTRCPADHPCPQWQJGEW 917

DB 841 WVIIEWGSCSKCELGWRRVLECRDINGQPASCAKEVKPASTRCPADHPCPQWQJGEW 900

QY 918 SSCSKTCKGKYKRSLSKLSHDSGVLSHESCDPLKKPKHPIDFCTMAECS 967

DB 901 SSCSKTCKGKYKRSLSKLSHDSGVLSHESCDPLKKPKHPIDFCTMAECS 950

Search completed: August 22, 2005, 10:15:39
Job time : 176 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	5232	99.0	967	4	US-09-130-491-2	Sequence 2, Appli
2	5140	97.2	949	4	US-09-568-559-2	Sequence 2, Appli
3	4260.5	80.6	950	4	US-09-321-987B-4	Sequence 4, Appli
4	3922	74.2	727	4	US-09-445-023A-1	Sequence 1, Appli
5	3656	69.2	727	4	US-09-445-023A-12	Sequence 12, Appli
6	3297	62.4	608	4	US-09-130-491-13	Sequence 13, Appli
7	2794	52.8	551	4	US-09-130-491-16	Sequence 16, Appli
8	2471.5	46.7	950	4	US-10-009-332-1	Sequence 1, Appli
9	2265.5	42.9	905	3	US-09-369-364A-9	Sequence 9, Appli
10	2124	40.2	837	4	US-09-122-126B-2	Sequence 2, Appli
11	2124	40.2	837	4	US-09-634-286A-2	Sequence 2, Appli
12	2124	40.2	837	4	US-10-247-685-2	Sequence 2, Appli
13	1924	36.4	930	4	US-09-122-126B-15	Sequence 15, Appli
14	1924	36.4	930	4	US-09-634-286A-15	Sequence 15, Appli
15	1924	36.4	930	4	US-10-247-685-15	Sequence 15, Appli
16	1309.5	36.1	930	3	US-09-369-364A-2	Sequence 2, Appli
17	1731.5	32.8	1892	3	US-09-369-364A-13	Sequence 13, Appli
18	1608.5	30.4	2150	4	US-09-321-987B-2	Sequence 2, Appli
19	1605.5	30.4	874	3	US-09-369-364A-15	Sequence 15, Appli
20	1598	30.2	2165	4	US-09-800-729-155	Sequence 155, Appli
21	1390.5	26.3	481	4	US-09-130-491-8	Sequence 8, Appli
22	1293	24.5	518	3	US-09-369-364A-22	Sequence 22, Appli
23	1287	24.3	1104	4	US-09-981-953A-4	Sequence 4, Appli
24	1278.5	24.2	1224	4	US-09-930-872-4	Sequence 4, Appli
25	1278.5	24.2	1224	4	US-10-217-774-4	Sequence 4, Appli
26	1214.5	23.0	908	4	US-09-963-791-2	Sequence 2, Appli
27	1207.5	22.8	937	3	US-09-369-364A-7	Sequence 7, Appli

QY 361 LCGSOTCDTLGMADYCTVCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKQASL 420
 DB 361 LCGSOTCDTLGMADYCTVCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKQASL 420
 QY 421 NGVNQDSHMMASMLNLDHSPWSPSCGYMITSFLDNGHGECLMDKPNQPIQLPGDLPGT 480
 DB 421 NGVNQDSHMMASMLNLDHSPWSPSCGYMITSFLDNGHGECLMDKPNQPIQLPGDLPGT 480
 QY 481 SYDANROCOFTPGEDSKHCPCDAASTCTLWCTGTSGGVLVCOQTKHPPWADGTCSEGGKWC 540
 DB 481 SYDANROCOFTPGEDSKHCPCDAASTCTLWCTGTSGGVLVCOQTKHPPWADGTCSEGGKWC 540
 QY 541 INKCKVKNHRKHFDTPHSGMWGMPGDCSRTCGGVQVYTMRECDNPVKNKGKCYCEG 600
 DB 541 INKCKVKNHRKHFDTPHSGMWGMPGDCSRTCGGVQVYTMRECDNPVKNKGKCYCEG 600
 QY 601 KRVYRSCNLEDCPDNNGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKL 660
 DB 601 KRVYRSCNLEDCPDNNGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKL 660
 QY 661 ICQAKGIGYFFVLQPKVVDGTPCSPDSTSVCGQCVKAGCDRIIDSXKXFKDKGVCVCGN 720
 DB 661 ICQAKGIGYFFVLQPKVVDGTPCSPDSTSVCGQCVKAGCDRIIDSXKXFKDKGVCVCGN 720
 QY 721 GSTCKKISGSVTSAPGVDHDIITPTGATNIEVKORNORSGRNNGSFLAIKAAADGTYILN 780
 DB 721 GSTCKKISGSVTSAPGVDHDIITPTGATNIEVKORNORSGRNNGSFLAIKAAADGTYILN 780
 QY 781 GDTLSTLEQDLMYKGVVLRYSGSSAALERSFSPLEPLTIQVLTGNALRPKIKYTY 840
 DB 781 GDTLSTLEQDLMYKGVVLRYSGSSAALERSFSPLEPLTIQVLTGNALRPKIKYTY 840
 QY 841 FVKKKESFNAIPTPSAWVIBEWGCSKSCBLGWORRLVECRDINGQPASECAKEVKRPAS 900
 DB 841 FVKKKESFNAIPTPSAWVIBEWGCSKSCBLGWORRLVECRDINGQPASECAKEVKRPAS 900
 QY 901 TRPCADHPCPQWLGEWSSCKTCGKYKTSKCLSHDGGVLSDHSCDPLKPKHFDIDF 960
 DB 901 TRPCADHPCPQWLGEWSSCKTCGKYKTSKCLSHDGGVLSDHSCDPLKPKHFDIDF 960
 QY 961 CTMAECS 967
 DB 961 CTMAECS 967

RESULT 2
 US-09-568-559-2
 ; Sequence 2, Application US/09568559
 ; Patent No. 6649377
 ; GENERAL INFORMATION:
 ; APPLICANT: Klonowski, Paul
 ; APPLICANT: Allard, John
 ; APPLICANT: Heller, Renu
 ; APPLICANT: Van Wart, Harold
 ; TITLE OF INVENTION: Human Aggrecanase and Nucleic Acid
 ; TITLE OF INVENTION: Compositions Encoding the Same
 ; FILE REFERENCE: ROCH-002
 ; CURRENT APPLICATION NUMBER: US/09/568,559
 ; CURRENT FILING DATE: 2000-05-09
 ; PRIOR APPLICATION NUMBER: 60/133,343
 ; PRIOR FILING DATE: 1999-05-10
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 949
 ; TYPE: PRT
 ; ORGANISM: human
 US-09-568-559-2
 Query Match 97.2%; Score 5140; DB 4; Length 949;
 Best Local Similarity 98.9%; Pred. No. 0;
 Matches 939; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 19 GNAERAPCSRSFGVPTLLILLAAALLAVSDALGRPSBEDELAVPELVRPVGHTTTRL 78
 DB 1 GNAERAPCSRSFGVPTLLILLAAALLAVSDALGRPSBEDELAVPELVRPVGHTTTRL 60
 QY 79 HAFQOQLDLDVPPDSSFLAPGFTLQNVGRKSGSTPLPETDLAHCYFSGTVNGPSSAAA 138
 DB 61 HAFQOQLDLDVPPDSSFLAPGFTLQNVGRKSGSTPLPETDLAHCYFSGTVNGPSSAAA 120
 QY 139 LSLCEGVRGAPYLLGEAYFIQPLPAASERLATAAPGKPPAPLOFHLRLRRNQDVGTC 198
 DB 121 LSLCEGVRGAPYLLGEAYFIQPLPAASERLATAAPGKPPAPLOFHLRLRRNQDVGTC 180
 QY 199 GVWDEPRPTKATEDEDETEGEDGEPWSPDPAALQGVQPTGCTSRKKBKPVSSHR 258
 DB 181 GVWDEPRPTKATEDEDETEGEDGEPWSPDPAALQGVQPTGCTSRKKBKPVSSHR 240
 QY 259 YVETMLVADQSMABFHGSGLKHYLLTLFSAARLYKHPSIRNSVSLVVKILVHDEQKG 318
 DB 241 YVETMLVADQSMABFHGSGLKHYLLTLFSAARLYKHPSIRNSVSLVVKILVHDEQKG 300
 QY 319 PEVTSNAALTILNFCNWKOHNPSPDRDAEHYDTAILFTRODLCGSQTCDTLGMADVGT 378
 DB 301 PEVTSNAALTILNFCNWKOHNPSPDRDAEHYDTAILFTRODLCGSQTCDTLGMADVGT 360
 QY 379 CDPSSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKQASLNGVNQDSHMMASMLNLD 438
 DB 361 CDPSSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKQASLNGVNQDSHMMASMLNLD 420
 QY 439 HSQPSWSPSCGYMITSFLDNGHGECLMDKPNQPIQLPGDLPGTSDANRQCOFTGEDSKH 498
 DB 421 HSQPSWSPSCGYMITSFLDNGHGECLMDKPNQPIQLPGDLPGTSDANRQCOFTGEDSKH 480
 QY 499 CPDAASTCTLWCTGTSGGVLVCOQTKHPPWADGTCSEGGKWCINGKCNKXKHFDP 558
 DB 481 CPDAASTCTLWCTGTSGGVLVCOQTKHPPWADGTCSEGGKWCINGKCNKXKHFDP 540
 QY 559 HSGMWGMPGDCSRTCGGVQVYTMRECDNPVKNKGKCYCEGKRVYRSCNLEDCPDNNG 618
 DB 541 HSGMWGMPGDCSRTCGGVQVYTMRECDNPVKNKGKCYCEGKRVYRSCNLEDCPDNNG 600
 QY 619 KTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLICQAKGIGYFFVLQPKV 678
 DB 601 KTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLICQAKGIGYFFVLQPKV 660
 QY 679 DGTPCSDDSTSVCGQCVKAGCDRIIDSXKXFKDKGVCVCGNSTCKKISGSVTSAPGY 738
 DB 661 DGTPCSDDSTSVCGQCVKAGCDRIIDSXKXFKDKGVCVCGNSTCKKISGSVTSAPGY 720
 QY 739 HDIITPTGATNIEVKORNORSGRNNGSFLAIKAAADGTYILNGDYTLSTLEQDLMYKGV 798
 DB 721 HDIITPTGATNIEVKORNORSGRNNGSFLAIKAAADGTYILNGDYTLSTLEQDLMYKGV 780
 QY 799 LRYSGSSAALERSFSPLEPLTIQVLTGNALRPKIKYTYFVKKKESFNAIPTPSAW 858
 DB 781 LRYSGSSAALERSFSPLEPLTIQVLTGNALRPKIKYTYFVKKKESFNAIPTPSAW 840
 QY 859 VIBEWGCSKSCBLGWORRLVECRDINGQPASECAKEVKRPASPCADHPCPQWLGEWS 918
 DB 841 VIBEWGCSKSCBLGWORRLVECRDINGQPASECAKEVKRPASPCADHPCPQWLGEWS 900
 QY 919 SCSTCKGKYKTSKCLSHDGGVLSDHSCDPLKPKHFDIDFCTMAECS 967
 DB 901 SCSTCKGKYKTSKCLSHDGGVLSDHSCDPLKPKHFDIDFCTMAECS 949

RESULT 3
 US-09-321-987B-4
 ; Sequence 4, Application US/09321987B
 ; Patent No. 6730820
 ; GENERAL INFORMATION:
 ; APPLICANT: Kimble, Judith E
 ; APPLICANT: Bielloch, Robert H

; TITLE OF INVENTION: Agent and Method for Modulating Cell Migration

; FILE REFERENCE: 960296.95386
; CURRENT APPLICATION NUMBER: US/09/321,987B
; CURRENT FILING DATE: 1999-05-28
; PRIOR APPLICATION NUMBER: 60/087,170
; PRIOR FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 60/129,023
; PRIOR FILING DATE: 1999-04-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 950
; TYPE: PRT
; ORGANISM: Murine
US-09-321-987B-4

Query Match 80.6%; Score 4260.5; DB 4; Length 950;
Best Local Similarity 81.6%; Pred. No. 0;
Matches 781; Conservative 56; Mismatches 105; Indels 15; Gaps 5;

QY 18 MGNABRAGRSRFGVPVPTLLLLAAA---LLAVSDALGRPSEDEELVPELSEVFGHG-- 72
DB 1 MGDVQBARSRGSLSAHMLLLLLASITMLLCARGAHRTEDEELVPSLERAFGHDS 60
QY 73 TTRLRLHAFDQOLDLDPDSSFLAPGFTLQNVGRKSGSDTPL--PETDLAHCFTYSGTVN 130
DB 61 TTRLRLDAFGQQLHLKLQDPSGLFAPGFTLQVGRSPGSEAQHLDPDPTGLAHCFTYSGTVN 120
QY 131 GDPSSAAALSCEGVRGAFYLLGEAYFTPLP-AAASERLATAAPGEKPPAPLOFHLLARN 189
DB 121 GDPGSAALSLCEGVRGAPYLLQGEFFIQAPGVATERLAPVPEESSARPPQHILARR 180
QY 190 ROGDVGGTCGVVDDPRPTGKAETDEDEGTGEDEGPQSPDPAQGVGQPTGTGSR 249
DB 181 RGSGGAKGVNDDELPTI-----SDSRPSQNRQMPVRDPTPDAGKSPGSGSR 233
QY 250 KKRFSVSHRYVETMLVADQSMAEFHGSLKHVLLTLFSVAARLYKHPSIRNSVLVVKI 309
DB 234 KKRFSVSPRYVETMLVADQSMADFHGSLKHVLLTLFSVAARFYKHPSIRNSISLVVKI 293
QY 310 LVIHDEQKGPVTSNAALTNRFCNWKQHNPPSDRDAEHYDTAILFTRODLCSGQTC 369
DB 294 LVIYEQKGPVTSNAALTNRFCNWKQHNPSDRDPEHYDTAILFTRODLCSGHTC 353
QY 370 LGMADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKQACASLNGVNQDSHM 429
DB 354 LGMADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKXCHASLNGVTDGSHL 413
QY 430 MASMLSNLDHSPWSPSCGYMITSFLDNGHGECMLMDKPNPIQLPGDLPGTSYDANRQC 489
DB 414 MASMLSSLDHSPWSPSCSAYMYVTSFLDNGHGECMLMDKPNPIKLPDLPGLTYDANRQC 473
QY 490 FTFGEDSKHCPDAASTCTLWCTGTSGGVLVCOQKHFPWADGTSCEGKWCINGKVCVN 549
DB 474 FTFGESKHCPCDAASTCTLWCTGTSGGVLVCOQKHFPWADGTSCEGKWCYSGKVCNKT 533
QY 550 HKRHKFDTPFHGSGMWGMPGDCSRCTCGGVQVYTMRECDNPVPKNGGKYCEGKRVYRSCN 609
DB 534 DMKHATPVHSGWGPMPGDCSRCTCGGVQVYTMRECDNPVPKNGGKYCEGKRVYRSCN 593
QY 610 LEDCPDNNKGTTFREEQCEAHNFEFSKASFGSGPAVEWIPKYAGVSPKDRCKLICQAKGIGY 669
DB 594 IEDCPDNNKGTTFREEQCEAHNFEFSKASFGNEPTVEWTPKYAGVSPKDRCKLTCEAKGIGY 653
QY 670 FVLQPKVVDGTPCSPDSTSVCVQGCYKAGCDRIIDSKKKPKDCKGVCNGNSTCKKISG 729
DB 654 FVLQPKVVDGTPCSPDSTSVCVQGCYKAGCDRIIDSKKKPKDCKGVCNGNSTCKKWSG 713
QY 730 SVTSKAPGVHDIITPTGATNTEVKORNGRNNGSFLAIIKAAQGTYLINGDYTLSTILE 789
DB 714 IVTSTRPGVHDIITPTGATNTEVKHNRNGRNNGSFLAIIRAADGTYILINGNFTLSTILE 773
QY 790 ODIMYKGVVRLYSGSAAALERSFSPLKEPLTIQVLTVGNALRPKIKYTFVFKKKKESF 849

DB 774 QDLTYKGTVRLYSGSSAALERSFSPLKBEPLTIQVLMVGHALRPKIKFTYFMKKKTSF 833
QY 850 NAIPTFSAWIEEWGECSKSCELGWQRRLVRECRDINGOPASECAKEVKPASTRCPCADHPC 909
DB 834 NAIPTFSEWIEEWGECSKTCGSGNQRRVVQCRDINGHPASECAKEVKPASTRCPCADLPC 893
QY 910 POWQLGEWSSCKTCGKGYKXKTSKLCSHSDGGVLSHSDCDPLKPKPKHFDICTMAEC 966
DB 894 PHWQVGDWSPCKTCGKGYKXKRTLKCVSHDGGVLSNESCDPLKPKPKHYIDFCTLTQC 950

RESULT 4

US-09-445-023A-1
; Sequence 1, Application US/09445023A
; Patent No. 6565858
; GENERAL INFORMATION:
; APPLICANT: Hiroee, Kunitaka
; APPLICANT: Inoguchi, Eiichi
; APPLICANT: Hakozaaki, Michinori
; APPLICANT: Ishioaka, Keiko
; APPLICANT: Ishida, Yukako
; APPLICANT: Matsuhashima, Kouji
; APPLICANT: Kuno, Kouji
; TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same, pharmaceutical
; FILE OF INVENTION: Composition and method of immunologically analyzing human ADAMTS-1
; FILE REFERENCE: Q57092
; CURRENT APPLICATION NUMBER: US/09/445,023A
; CURRENT FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: JP 9-160422
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent In version 3.0
; SEQ ID NO 1
; LENGTH: 727
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-445-023A-1

Query Match 74.2%; Score 3922; DB 4; Length 727;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 707; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 248 IRKRFVSSHRYVETMLVADQSMAEFHGSLKHVLLTLFSVAARLYKHPSIRNSVLVWV 307
DB 8 LRKRFVSSPRYVETMLVADQSMAEFHGSLKHVLLTLFSVAARLYKHPSIRNSVLVWV 67
QY 308 KILVIHDEQKGPVTSNAALTNRFCNWKQHNPPSDRDAEHYDTAILFTRODLCSGQTC 367
DB 68 KILVIHDEQKGPVTSNAALTNRFCNWKQHNPPSDRDAEHYDTAILFTRODLCSGQTC 127
QY 368 DTGLMADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKQACASLNGVNQDS 427
DB 128 DTGLMADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKQACASLNGVNQDS 187
QY 428 HMAMSLNLDHSPWSPSCGYMITSFLDNGHGECMLMDKPNPIQLPGDLPGTSYDANRQ 487
DB 188 HMAMSLNLDHSPWSPSCSAYMITSFLDNGHGECMLMDKPNPIQLPGDLPGLTYDANRQ 247
QY 488 COFTFGEDSKHCPDAASTCTLWCTGTSGGVLVCOQKHFPWADGTSCEGKWCINGKVCVN 547
DB 248 COFTFGEDSKHCPDAASTCTLWCTGTSGGVLVCOQKHFPWADGTSCEGKWCINGKVCVN 307
QY 548 KNHRKHFDTPFHGSGMWGMPGDCSRCTCGGVQVYTMRECDNPVPKNGGKYCEGKRVYRSCN 607
DB 308 KTDHRKHFDTPFHGSGWGPMPGDCSRCTCGGVQVYTMRECDNPVPKNGGKYCEGKRVYRSCN 367
QY 608 CNLEDCPDNNKGTTFREEQCEAHNFEFSKASFGSGPAVEWIPKYAGVSPKDRCKLICQAKGI 667
DB 368 CNLEDCPDNNKGTTFREEQCEAHNFEFSKASFGSGPAVEWIPKYAGVSPKDRCKLICQAKGI 427
QY 668 GYFFVLQPKVVDGTPCSPDSTSVCVQGCYKAGCDRIIDSKKKPKDCKGVCNGNSTCKKI 727

```
Db 428 GYFFVLQPKVVDGTPCSPDSTSVCGQCVKAGCDRIIDSKKKFKDKCGVCGGNGSTCKKI 487
Qy 728 SGSVTSAPGYHDIITPTGATNIEVKQNRGRNNGSFLAIKAADGTYILNGDYTLST 787
Db 488 SGSVTSAPGYHDIITPTGATNIEVKQNRGRNNGSFLAIKAADGTYILNGDYTLST 547
Qy 788 LEQDIMYKGVILRYSGSSAALRIIRSFSPKLEPLTIQVLTGVNALRPKIKYTYFVKKKKE 847
Db 548 LEQDIMYKGVILRYSGSSAALRIIRSFSPKLEPLTIQVLTGVNALRPKIKYTYFVKKKKE 607
Qy 848 SFNAIPTSASWIEBWGECSKCELGWQRLVECRDINGOPASECAKEVKPASTRPCADH 907
Db 608 SFNAIPTSASWIEBWGECSKCELGWQRLVECRDINGOPASECAKEVKPASTRPCADH 667
Qy 908 PCPQOLGWSKSTCGKGYKKTSLKCLSHDGGVLSHSDCDPLKKPKHFIDFCTMAECS 967
Db 668 PCPQOLGWSKSTCGKGYKKTSLKCLSHDGGVLSHSDCDPLKKPKHFIDFCTLTQCS 727

RESULT 5
US-09-445-023A-12
; Sequence 12, Application US/09445023A
; Patent No. 6565858
; GENERAL INFORMATION:
; APPLICANT: Hirose, Kunitaka
; APPLICANT: Inoguchi, Eiichi
; APPLICANT: Hakozaaki, Michinori
; APPLICANT: Ishioka, Keiko
; APPLICANT: Ishida, Yukako
; APPLICANT: Matsushima, Kouji
; APPLICANT: Kuno, Kouji
; TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same, pharmaceutical
; FILE REFERENCE: Q57092
; CURRENT APPLICATION NUMBER: US/09/445,023A
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: JP 9-160422
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 727
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-445-023A-12

Query Match 69.2%; Score 3656; DB 4; Length 727;
Best Local Similarity 89.0%; Pred. No. 1.9e-311;
Matches 647; Conservative 42; Mismatches 38; Indels 0; Gaps 0;

Qy 241 OPTGTGIRKRRFYSSHYRVTMLVADQSMAEFHGSGLKHVLLTLFSVAARLYKHPSTRN 300
Db 1 EPSGPGIRKRRFYSSHYRVTMLVADQSMADFHGSGLKHVLLTLFSVAARLYKHPSTRN 60
Qy 301 SVSLVVKVILVIHDEQKGPVTSNAALTIRNFCNWKQKHNPPSDRDAEHYDTAILFTTRQD 360
Db 61 SISLVVVKVILVIHDEQKGPVTSNAALTIRNFCNWKQKHNPPSDRDAEHYDTAILFTTRQD 120
Qy 361 LCGSQTCDTLGMADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACSL 420
Db 121 LCGSHTCDTLGMADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKHCA 180
Qy 421 NGVNQDSHMASMLSLNDHSPWSPGCSGYMTSFLDNGHGECCLMDKPNPIQLPGDLPGT 480
Db 181 NGVSDSHMASMLSLNDHSPWSPGCSGYMTSFLDNGHGECCLMDKPNPIQLPGDLPGT 240
Qy 481 SYDANRQCQFTFGEDSKHCPDAASTCTSLWCTGTSGGVLCVQTKHFPWADGTSCEGKWC 540
Db 241 LYDANRQCQFTFGEDSKHCPDAASTCTSLWCTGTSGGVLCVQTKHFPWADGTSCEGKWC 300
Qy 541 INKCVNKHRRKHEDTTPHGSWGMWGPWGDGCSRTCGGGVQYTMRECDNPPKNGGKYCEG 600
Db 301 VSGKCVNKTDMKHFATPVHGSWGMWGPWGDGCSRTCGGGVQYTMRECDNPPKNGGKYCEG 360
```

```
Qy 601 KRVRYSNLEDCPDNNGKTFREBQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKL 660
Db 361 KRVRYSNLEDCPDNNGKTFREBQCEAHNEFSKASFGNEPTWTPKYAGVSPKDRCKL 420
Qy 661 ICOAKGIGYFFVLQPKVVDGTPCSPDSTSVCGQCVKAGCDRIIDSKKKFKDKCGVCGGN 720
Db 421 TCEAKGIGYFFVLQPKVVDGTPCSPDSTSVCGQCVKAGCDRIIDSKKKFKDKCGVCGGN 480
Qy 721 GSTCKKISGVSATSAKPGYHDIITPTGATNIEVKQNRGRNNGSFLAIKAADGTYILN 780
Db 481 GSTCKKISGIVTSTRPGYHDIITPTGATNIEVKQNRGRNNGSFLAIKAADGTYILN 540
Qy 781 GDYTLSTLEQDIMYKGVILRYSGSSAALRIIRSFSPKLEPLTIQVLTGVNALRPKIKYTY 840
Db 541 GNFTLSTLEQDIMYKGVILRYSGSSAALRIIRSFSPKLEPLTIQVLMVGHALRPKIKYTY 600
Qy 841 FVKKKESFNAIPTFSAWIEBWGECSKCELGWQRLVECRDINGOPASECAKEVKPAS 900
Db 601 FMKKKTESFNAIPTFSAWIEBWGECSKCELGWQRLVECRDINGOPASECAKEVKPAS 660
Qy 901 TRPCADHPCPQOLGWSKSTCGKGYKKTSLKCLSHDGGVLSHSDCDPLKKPKHFIDF 960
Db 661 TRPCADLPCHWQVGDWSPCSTCGKGYKKTSLKCLSHDGGVLSHSDCDPLKKPKHYIDF 720
Qy 961 CTMAECS 967
Db 721 CTLTQCS 727

RESULT 6
US-09-130-491-13
; Sequence 13, Application US/09130491
; Patent No. 6416974
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
; FILE REFERENCE: 09404/041001
; CURRENT APPLICATION NUMBER: US/09/130,491
; CURRENT FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: US 60/058,108
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: US 60/054,961
; EARLIER FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 608
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-130-491-13

Query Match 62.4%; Score 3297; DB 4; Length 608;
Best Local Similarity 90.9%; Pred. No. 4.3e-280;
Matches 599; Conservative 1; Mismatches 5; Indels 54; Gaps 1;
```

```
Qy 309 ILVIHDEQKGPVTSNAALTIRNFCNWKQKHNPPSDRDAEHYDTAILFTTRQDLCGSQTC 368
Db 4 ILVIHDEQKGPVTSNAALTIRNFCNWKQKHNPPSDRDAEHYDTAILFTTRQDLCGSQTC 63
Qy 369 TLGMADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACSLANGVNQDSH 428
Db 64 TLGMADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACSLANGVNQDSH 123
Qy 429 MMASMLSLNDHSPWSPGCSGYMTSFLDNGHGECCLMDKPNPIQLPGDLPGTSDYDANRQC 488
Db 124 MMASMLSLNDHSPWSPGCSGYMTSFLDNGHGECCLMDKPNPIQLPGDLPGTSDYDANRQC 183
Qy 489 QFTFGEDSKHCPDAASTCTSLWCTGTSGGVLCVQTKHFPWADGTSCEGKWCINGKCVNK 548
Db 184 QFTFGEDSKHCPDAASTCTSLWCTGTSGGVLCVQTKHFPWADGTSCEGKWCINGKCVNK 243
```

```

QY 549 NHRKHFDTPFHSGWGMWGPWGDSCRTCCGGVQVYTMRECDNPVPKNGKGYCEGKRVYRSC 608
DB 244 TDRKHFDTPFHSGWGMWGPWGDSCRTCCGGVQVYTMRECDNPVPKNGKGYCEGKRVYRSC 303
QY 609 NLEDCPDNNGKTFREOCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLICQAKGIG 668
DB 304 NLEDCPDNNGKTFREOCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLICQAKGIG 363
QY 669 YFVLQPKVVDGTPCSPDSTSVCCVQCVKAGCDRIIDSKKKFDKCGVCGGNGSTCKKIS 728
DB 364 YFVLQPKVVDGTPCSPDSTSVCCVQCVKAGCDRIIDSKKKFDKCGVCGGNGSTCKKIS 423
QY 729 GSVTSKAPGYHDIITPTGATNIEVKQNRQSRNNGSFLAIKAADGTYLNGDYTLSTL 788
DB 424 GSVTSKAPGYHDIITPTGATNIEVKQNRQSRNNGSFLAIKAADGTYLNGDYTLSTL 483
QY 789 EODIMYGVVLYRSGSSAALERIRSFPLKEPTIQVLTGVNLRPKIKYTFVKKKES 848
DB 484 EODIMYGVVLYRSGSSAALERIRSFPLKEPTIQVLTGVNLRPKIKYTFVKKKES 543
QY 849 FNAIPTFSAWIEEWGECSEKSCELGWQRRLVECRDINGQPASECAKEVKPASTRPCADHP 908
DB 544 FNAIPTFSAWIEEWGECSEKSCELGWQRRLVECRDINGQPASECAKEVKPASTRPCADHP 560
QY 909 CPWQLGEWSSCKTCGKGYKTKSLKCLSHDGGVLSHSDCDPLKPKPHFIDFCTMAECS 967
DB 561 -----CSKTCGKGYKTKSLKCLSHDGGVLSHSDCDPLKPKPHFIDFCTMAECS 608

RESULT 7
US-09-130-491-16
; Sequence 16, Application US/09130491
; Patent No. 6416974
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
; FILE REFERENCE: 09404/041001
; CURRENT APPLICATION NUMBER: US/09/130,491
; CURRENT FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: US 60/058,108
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: US 60/054,961
; EARLIER FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 551
; TYPE: PRT
; ORGANISM: Rattus rattus
US-09-130-491-16

```

```

Query Match 52.8%; Score 2794; DB 4; Length 551;
Best Local Similarity 88.0%; Pred. No. 4.7e-236;
Matches 485; Conservative 35; Mismatches 31; Indels 0; Gaps 0;

QY 417 CASLNGVNDSHMASMLNLDHSPQSPSCGYMTSFLDNGHGECCLMDKPNQIQLPGD 476
DB 1 CASLNGVSDSHLMASMLSLDHSQWPSPCSAYMTSFLDNGHGECCLMDKPNQIQLPSD 60
QY 477 LPTGTSYDANRQCOFTGEDSKHCPDAASTCLTWCCTGTSGGVLVCQTKHFPWADGTSCE 536
DB 61 LPTGLYDANRQCOFTGSESKHCPDAASTCLTWCCTGTSGLLVCTKHFPWADGTSCE 120
QY 537 GWKCNKGVKNKHXHPTPHGSGMGPWGDSCRTCCGGVQVYTMRECDNPVPKNGK 596
DB 121 GWKCVSGKCNKHTDMKHFAIPVHSGMGPWGDSCRTCCGGVQVYTMRECDNPVPKNGK 180
QY 597 YCEGKRVYRSCNLEDCPDNNGKTFREOCEAHNEFSKASFGSGPAVEWIPKYAGVSPKD 656
DB 181 YCEGKRVYRSCNLEDCPDNNGKTFREOCEAHNEFSKASFGNEPTVETPKYAGVSPKD 240
QY 657 RCKLICQAKGIGYFFVLQPKVVDGTPCSPDSTSVCCVQCVKAGCDRIIDSKKKFDKCGV 716

```

```

DB 241 RCKLTCEAKGIGYFFVLQPKVVDGTPCSPDSTSVCCVQCVKAGCDRIIDSKKKFDKCGV 300
QY 717 CGNGSTCKKISGVSYSKAPGVYHDIITPTGATNIEVKQNRQSRNNGSFLAIKAADGT 776
DB 301 CGNGSTCKKISGVSYSKAPGVYHDIITPTGATNIEVKQNRQSRNNGSFLAIKAADGT 360
QY 777 YILNGDYTLSTLEQDQIMYKGVVLYRSGSSAALERIRSFPLKEPTIQVLTGVNLRPKI 836
DB 361 YILNGDYTLSTLEQDQIMYKGVVLYRSGSSAALERIRSFPLKEPTIQVLTGVNLRPKI 420
QY 837 KYTYFVKKKESFNAIPTFSAWIEEWGECSEKSCELGHQRRLVECRDINGQPASECAKEV 896
DB 421 KYTYFVKKKESFNAIPTFSAWIEEWGECSEKSCELGHQRRLVECRDINGQPASECAKEV 480
QY 897 KPASTRPCADHPCPWQLGEWSSCKTCGKGYKTKSLKCLSHDGGVLSHSDCDPLKPKPH 956
DB 481 KPASTRPCADHPCPWQLGEWSSCKTCGKGYKTKSLKCLSHDGGVLSHSDCDPLKPKPH 540
QY 957 FIDFCTMAECS 967
DB 541 VIDFCTLTQCS 551

RESULT 8
US-10-009-332-1
; Sequence 1, Application US/10009332
; Patent No. 6716613
; GENERAL INFORMATION:
; APPLICANT: Yamanouchi Pharmaceutical Co., Ltd.
; APPLICANT: Kazusa DNA Research Institute
; TITLE OF INVENTION: NOVEL METALLOPROTEASE HAVING AGGREGANASE ACTIVITY
; FILE REFERENCE: Q67541
; CURRENT APPLICATION NUMBER: US/10/009,332
; CURRENT FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: JPA Hei 11-321740
; PRIOR FILING DATE: 1999-11-11
; PRIOR APPLICATION NUMBER: JPA 2000-144020
; PRIOR FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 950
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-009-332-1

```

```

Query Match 46.7%; Score 2471.5; DB 4; Length 950;
Best Local Similarity 48.8%; Pred. No. 2.1e-207;
Matches 484; Conservative 153; Mismatches 253; Indels 101; Gaps 24;

QY 36 LLLLAALLAVSDALGRPSEDEELVVP-ELE-----RVP-----GHGTTTRLRLHAF 81
DB 1 MLLGILTLAFAGTAGGFEREVVFIRLDPDINGRRYWRGPEDSGDGLIFQITAF 60
QY 82 DQDLDDVPDPSSFLAPGFTLLQNVGRKSGSDTLP-----ETDLAHCFYSYGVNGDPSSA 136
DB 61 QEDFYHLTLPDAQFLAPAFSTEHLG-----VPLQGLTGSSDLRRRCFYSGDVNAEPDSF 114
QY 137 AALSICEGVGAFYLLGEAYFIQPLPAASERLATAPGEKPPAPLOPHLLRRNRQDVG 196
DB 115 AAVSLCGGLRGAFGYRGAEYVISPLPNAS---APAAQRNSQGA---HLLQ---RRGVFG 165
QY 197 TGVVVDDEPRTGKAEDEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGE 247
DB 166 PSG-----DPTSRC-----GVASG---WNPAILRALDPYKPRAGFGESRRRS 207
QY 248 IRKKRFVSHRYVETMLVADQSMAEFGHSGKHLHYLLTLFVSAARLYKHPSIRNSVSLVV 307
DB 208 GRAKEFVSI PRVETLVVVADESVMKFGHADLEHYLLTLTAARLYRHPSILNPINIVV 267
QY 308 KILVTHDSQKPEVTSNAALTIRNFCNWKQKHNPSPDRDAEHYDTAILFTTRQDLCSQTC 367

```


; NUMBER OF SEQ ID NOS: 21
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 837
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-122-126B-2

Query Match 40.2%; Score 2124; DB 4; Length 837;
 Best Local Similarity 49.0%; Pred. No. 5e-177;
 Matches 411; Conservative 132; Mismatches 225; Indels 70; Gaps 18;

Qy	36	LLLLAAALLAVSDALGRPSEDEELVPEL-----ERVPGHGT-TRL--RLHAFDQQLDLVD	89
Db	37	LLLLLLASLLPSARLASPLPREEEIVFPEKLGSLVPGSGAPARLLCLRLQAFGETLLLEL	96
Qy	90	PPDSFLAPGFTLQNVGRK-----SGSDTLPETDLAHCFYSGTVNGDPSSAALSICSG-	144
Db	97	BODSGVQVEGLTVYLGQAPPELLGAE---PGT-----YLTGTINGDPESVASLHWDGA	148
Qy	145	VRGAFYLLGEAYFIQPLPAASERLATAAPGKPPAPLQFHLRLRRNQDVGCTCGVVDDE	204
Db	149	LLGVLYRGAEHLQPLEGGTNSA--GGPGA-----HILRRK-----	184
Qy	205	PRPTGKAETEDDEGTEGEDGEPQSPDPALQGVQGTGTSIRKKRPFVSHRYVETML	264
Db	185	-----SPASGQGPNCNVKAP-----LGSPSPRPR-RAKRFASLSRREVETLV	224
Qy	325	AALTLRNFQKQHNPPSDRDAEHYDTAILFTRODLGCSQTCDTLGMADVGTVCDPSPRS	384
Db	285	AAQTILRSFCANQGLNTPEDSDPHDFDTAILFTRODLGCVSTCDTLGMADVGTVCDPARS	344
Qy	385	CSVIEDGLQAAFTTAHELGHVFNPHDDAKQASLNG--VNQDSHMMASMLSNLDHSPW	443
Db	345	CAIVEDDLQSAFTAHAHELGHVFNMLHDNSKPCISLNGPLSTSRHVMAVMAHVDPPEPW	404
Qy	444	SPCSARFITDFLDNGYGHCLLDKPEAPLHLPVTFPGKDYDADROCOLTFGPDSPRHCPLP	464
Db	504	STCSTLWCTGTSGGVVLCQTKHFPWADGTSCEGKWCINGKCVNKNHRKHFDTPPHGSGW	563
Qy	624	EOCEAHNEFSKASFCSGPA--VEMI PKYAGVSPKDRCKLI COAKGIGYFVLQPKVVDGTP	682
Db	585	EQCAAYNHRTDL--FKSPFGPMDVPRYTGVAPODQCKLTQARALGYVYVLEPRVVDGTP	643
Qy	683	CSPDSTSVCGQCVKAGCDRIIDSKKFKDKCGVCGNGSTCKKISGVS TSAKPGYHDI	742
Db	644	CSPDSSSVCGQRCIHAGCDRIIGSKKFKDKCMVCGDGGSGCSKQSGFRFRYGYNNV	703
Qy	743	TIPTGATNIEVKQNRQGRNNGSFLAIKAADGTYYILNGDYTLSTLEODIMYKGVV--LRY	801
Db	704	TIPAGATHILVRQQGNPGHRS--IYLALKLPDGSYALNGEYTLMPSPDTPVLPAGVSLRY	761
Qy	802	SGSSAALERIRSFPLKEPLTIOVLTVGNALRPKIKYTYFVKKKKESNAIPTSAW	859
Db	762	SGATAASETLSGHGFLAQPLTLQVLVAGNPQDTRLRYSFVFRPTPS--TPRPTQDWL	818

RESULT 11

US-09-634-286A-2
 ; Sequence 2, Application US/09634286A
 ; Patent No. 6521436

; GENERAL INFORMATION:
 ; APPLICANT: Bristol-Myers Squibb Company
 ; TITLE OF INVENTION: AGGRECAN DEGRADING METALLO PROTEASES
 ; FILE REFERENCE: DM6909A
 ; CURRENT APPLICATION NUMBER: US/09/634.286A
 ; CURRENT FILING DATE: 2000-08-09
 ; NUMBER OF SEQ ID NOS: 21
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 837
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-634-286A-2

Qy	36	LLLLAAALLAVSDALGRPSEDEELVPEL-----ERVPGHGT-TRL--RLHAFDQQLDLVD	89
Db	37	LLLLLLASLLPSARLASPLPREEEIVFPEKLGSLVPGSGAPARLLCLRLQAFGETLLLEL	96
Qy	90	PPDSFLAPGFTLQNVGRK-----SGSDTLPETDLAHCFYSGTVNGDPSSAALSICSG-	144
Db	97	BODSGVQVEGLTVYLGQAPPELLGAE---PGT-----YLTGTINGDPESVASLHWDGA	148
Qy	145	VRGAFYLLGEAYFIQPLPAASERLATAAPGKPPAPLQFHLRLRRNQDVGCTCGVVDDE	204
Db	149	LLGVLYRGAEHLQPLEGGTNSA--GGPGA-----HILRRK-----	184
Qy	205	PRPTGKAETEDDEGTEGEDGEPQSPDPALQGVQGTGTSIRKKRPFVSHRYVETML	264
Db	185	-----SPASGQGPNCNVKAP-----LGSPSPRPR-RAKRFASLSRREVETLV	224
Qy	265	VADQSMARFHGSLKHYLLTLFSVAARLYKHPSIRNSVLSVVKLVLIHDEQKGPVTSN	324
Db	225	VADQMAAFHAGLAKRYLLTVMAAAKAFKPSIRNPVSLVTVRLVILGSGEEGQVGPVS	284
Qy	325	AALTLRNFQKQHNPPSDRDAEHYDTAILFTRODLGCSQTCDTLGMADVGTVCDPSPRS	384
Db	285	AAQTILRSFCANQGLNTPEDSDPHDFDTAILFTRODLGCVSTCDTLGMADVGTVCDPARS	344
Qy	385	CSVIEDGLQAAFTTAHELGHVFNPHDDAKQASLNG--VNQDSHMMASMLSNLDHSPW	443
Db	345	CAIVEDDLQSAFTAHAHELGHVFNMLHDNSKPCISLNGPLSTSRHVMAVMAHVDPPEPW	404
Qy	444	SPCSARFITDFLDNGYGHCLLDKPEAPLHLPVTFPGKDYDADROCOLTFGPDSPRHCPLP	464
Qy	504	STCSTLWCTGTSGGVVLCQTKHFPWADGTSCEGKWCINGKCVNKNHRKHFDTPPHGSGW	563
Db	465	PPCAALWCSGHLNGHAMCOTKHS PWADGTPCGPAQACMGRCCLHMDQLODFNI POAGG	524
Qy	564	MWGPWGDCSRTCGGVQVYTMRECDNPVPKNGKCYCEGKRVYRSCNLEDCPDNNGKTRE	623
Db	525	PWGPWGDCSRTCGGVQVFSRRDCTRPVPKNGKCYCEGKRVYRSCNLEDCPDNNGKTRE	584
Qy	624	EOCEAHNEFSKASFCSGPA--VEMI PKYAGVSPKDRCKLI COAKGIGYFVLQPKVVDGTP	682
Db	585	EQCAAYNHRTDL--FKSPFGPMDVPRYTGVAPODQCKLTQARALGYVYVLEPRVVDGTP	643
Qy	683	CSPDSTSVCGQCVKAGCDRIIDSKKFKDKCGVCGNGSTCKKISGVS TSAKPGYHDI	742
Db	644	CSPDSSSVCGQRCIHAGCDRIIGSKKFKDKCMVCGDGGSGCSKQSGFRFRYGYNNV	703
Qy	743	TIPTGATNIEVKQNRQGRNNGSFLAIKAADGTYYILNGDYTLSTLEODIMYKGVV--LRY	801
Db	704	TIPAGATHILVRQQGNPGHRS--IYLALKLPDGSYALNGEYTLMPSPDTPVLPAGVSLRY	761
Qy	802	SGSSAALERIRSFPLKEPLTIOVLTVGNALRPKIKYTYFVKKKKESNAIPTSAW	859
Db	762	SGATAASETLSGHGFLAQPLTLQVLVAGNPQDTRLRYSFVFRPTPS--TPRPTQDWL	818

RESULT 12

US-10-247-685-2
; Sequence 2, Application US/10247685
; Patent No. 6753176
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: AGGREGAN DEGRADING METALLO PROTEASES
; FILE REFERENCE: DM6909D
; CURRENT APPLICATION NUMBER: US/10/247,685
; CURRENT FILING DATE: 2002-09-19
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 837
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-247-685-2

Query Match 40.2%; Score 2124; DB 4; Length 837;
Best Local Similarity 49.0%; Pred. No. 5e-177;
Matches 411; Conservative 132; Mismatches 225; Indels 70; Gaps 18;
Vg
36 LLLLLAALLAVSDALGRPESEDEELVPEL---ERVPGHGT-TRL--RLHAFDQOLDLDV 89
Db
37 LLLLLLASLLPSARLASPLPREEEIVFPEKLSVLPFGSGAPARLLCLRLQAFGETLLEL 96
Qy 90 PPDSSFLAPGFTLQNVGRK---SGSDTLPETDLAHCFYGTVNGDPSSAAALSCLCEG- 144
Db 97 EQDSGVQVEGLTVQLGQAPPELLGAE---PGT-----VLTGTINGDPESVASLHWDGGA 148
Qy 145 VRGAPYLGEAYFIOLPAASERLATAAPGKPPAPLOFHLRLRRNRQDVGTCGVVDDE 204
Db 149 LIGVLQYRGAEHLPLQLEGGTPNSA--GGPGA-----HILRRK----- 184
Qy 205 PRPTKAEATEDEDETEGEDEGPQSPDPAALQGVGQPTGTSIRKRFVSSHRYVETML 264
Db 185 -----SPASGQGPNCNVKAP-----LGSPSPRP-RAKGFASLSRFVETLV 224
Qy 265 VADQSMABFHGSLKHLYLLTFSVAARLYKHPISIRNSVSVVVKILVTHDQKGEVTSN 324
Db 225 VADDKMAAFHAGLAKRYLLTVMAAAAKAFKHPISIRNPVSVLVVTRIVILGSGEGPQVGPS 284
Qy 325 AALTLRNCNWKQHNPSDRDAEHVDTAILFTRODLGSGTCDTLGMADVGTVCDDPSR 384
Db 285 AAQTLRSFCAMQRLNTPEDSDPHDFDTAILFTRODLGSGTCDTLGMADVGTVCDDPARS 344
Qy 385 CSVIEDDGLQAAFTTAHELGHVFNPNPHDDAKQASLNG-VNQDSHMMASMLSLNLDHSQPW 443
Db 345 CAIVEDDGLQSAFTAAHELGHVFNMLHNDNSKPCISLNGPLSTSRHVMAPVMAHVDPEEPW 404
Qy 444 SPCSGYMITSLDNGHGLMDKPNQPTOLPGDLPGTSDYDANRQCQFTFGSDSKHCPDAA 503
Db 405 SPCSAARFTDFLDNGYGHCLLDKPEAPLHLPTVTFPGDYDADROCOLTFGPDSDRHCPOLP 464
Qy 504 STCSLTWCTTSGGVLCVOTKHPHADGTSCEGKWCINGKCNVKNHKKHFDTPFHSGWG 563
Db 465 PPCAALWCSHLNGHAMCOTKXHSWADGTPCPGAQAQMGGRCLHWDQLQDNFIQAQGWG 524
Qy 564 MWGPWGDCSRCTCGGVQVYTRRECONVPKNGKCYCEGKRVYRSCLNLDGCDNNGKTFRE 623
Db 525 PWGPWGDCSRCTCGGVQFSRDRCTRPVPRNGKCYCEGRTFRSCNTEDCTGSAITFRE 584
Qy 624 EQCEAHNEFSKASFGSGPA-VIEWIPKYAGVSPKORCKLIQAKGIGYFVLQPKVVDGTP 682
Db 585 EQCAAYNHRIDL-FKSPFPGMDWVPRYTGVAPOQOQCKLTQCARALGYVYVLEPRVVDGTP 643
Qy 683 CSPDSTSCVQOCVKACCDRIIDSKKFDKCGVCGGNGSTCKKISGVSPTSACKEGYDII 742
Db 644 CSPDSSSVCVQGRCIHACCDRIIGSKKFDKCMVCGGNGSGCSKQSGFRFRKYGNVNV 703
Qy 743 TIPTGATNIEVKQRNQRNRNGSFLAIKAADGTGYILNGDYLTLSTLEQDIIYKGVV-LRY 801

Db 704 TIPAGATHILVRQGNPCHRS--IYLALKLPDGSYALNGEYTLMPSPDVLPGAVSLRY 761
Qy 802 SGSSAALERIRSFSPLEKPLTIQVLTGVGNALRPKIKYTYFVKKKKESNAIPTSANV 859
Db 762 SGATAASETTLGCHGLAQPLTLQVLVAGNPQDTRLRYSEFFVPRPTPS--TPRPTQDWL 818
RESULT 13
US-09-122-126B-15
; Sequence 15, Application US/09122126B
; Patent No. 6451575
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: AGGREGAN DEGRADING METALLO PROTEASES
; FILE REFERENCE: DM6909
; CURRENT APPLICATION NUMBER: US/09/122,126B
; CURRENT FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-122-126B-15
Query Match 36.4%; Score 1924; DB 4; Length 930;
Best Local Similarity 40.7%; Pred. No. 2e-159;
Matches 391; Conservative 125; Mismatches 314; Indels 130; Gaps 17;
Vg
25 PGRSFGPVPPTLLLLAALLAVSDALGRPESEDEELVPELERVPGHG---TTRLRLHAF 81
Db 26 PAQDKAQGPPT-----AAAAQPRRQGEVEVERAEP-----PGHPHLQAORRRSGL 73
Qy 82 DQQLD-----LDVPPDSSFLAPGFTLQNVGRKSGSDTLPPLPDL 120
Db 74 VNQDQLYSGGKGVLYVYAGRRFLDLERDGSVGIAGFV-----PAGGTSAPWRHR 127
Qy 121 AHCFYSGTVNGDPSSAAALSICEVRCGAFYLLGEAYFIQPL---PAASER----- 167
Db 128 SHCFYRGTVDASPRSLAVFDLCGLDGFVAVKHAARYTLKPLLRGPWABEEKRGVYGDGA 187
Qy 168 -----LATAAPGKPPAPLOFHLRLRRNRQDVGTCGVVDDEP---RPTG 209
Db 188 RILHVVYREGFSFPAALPPRASCETPASTPEAH-----EHAPAHSPSG 230
Qy 210 KA-----ETDEDETEGEDEGPQ--WSPQDPAALQGVGQPTGTSIRKRFVSSHRYVETM 263
Db 231 RAALASQLLDQSALSPAGSGSPQTMWR-----RRRSISRARQVELL 272
Qy 264 LVADQSMABFHGSLKHLYLLTFSVAARLYKHPISIRNSVSVVVKILVTHDQKGEVTS 323
Db 273 LVADASMARLYRGLOHLYLLTASIANRLYSHASIEHIRLAVVVKVVLVGDGDKSLEYSK 332
Qy 324 NAALTLRNCNWKQHNPSDRDAEHVDTAILFTRODLGSGTCDTLGMADVGTVCDDPSR 383
Db 333 NAATLKNFCKWQHQNQLGDDHEEHYDAAILFTREDLCGHSCDTLGMADVGTICSPER 392
Qy 384 SCSVIEDDGLQAAFTTAHELGHVFNPNPHDDAKQASLNGVQNQDSHMMASMLSLNLDHSQPW 443
Db 393 SCAVIEDDGLHAAFTVAHEIGHLLGLSHDSDSKFCEETFGSTEDKRLMSSILUTSIDAKPW 452
Qy 444 SPCSGYMITSLDNGHGLMDKPNQPTOLPGDLPGTSDYDANRQCQFTFGSDSKHCPDAA 503
Db 453 SKCTSATITBFLDDHGNCLLDLPRKQILGPEELPGQTYDATQOCNLTGPEYSVCP- 511
Qy 504 STCSLTWCTTSGGVLCVOTKHPHADGTSCEGKWCINGKCNVKNHKKHFDTPFHSGWG 563
Db 512 DVCARLWCAVVRQGMVCLTKLPAVBGTFCGKGRICLQCKVDKTKKKYVSTSSHNGW 571
Qy 564 MWGPWGDCSRCTCGGVQVYTRRECONVPKNGKCYCEGKRVYRSCLNLDGCDNNGKTFRE 623
Db 572 SWGSGQCSRSCGGGVQFAYRHCNPNRNGRYCTGKRAIYRSCSLMPCPP--NGKSF 630

QY 624 ECEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLICQAKGIGYFFVLQPKVVDGTPC 683
 DB 631 ECEAKNGYQSDAKGVKTEFVWPKYAGVLPADVCKLTCRAKGTGYVVFSPKVTDGTEC 690
 QY 684 SPDSTSVCOGCVKAGCDRIIDSKKPKFCGCGNGSTCKKISGVSATSAKPGVHDIIT 743
 DB 691 RPYNSVVRGKCVRTGCGIIGSKLQYDKGCGGDNSSCTKIIVTFNKKSKGYTDVVR 750
 QY 744 IPTGATNIEVKORNGSRNNGSFLAIKAADGTIYILNGDYTLSTLEQDLMYKGVVLYRSG 803
 DB 751 IPEGATHIKVRQKAKQDRTFTAYLALKKNGEYLINGKYMISTETIIDINGTVMNYS 810
 QY 804 SSAALERIR--SFSPKLEPLTIQVLTVGNALRPKIYTYFV-KKKESFNAIPTFSA--- 857
 DB 811 WSHRDDFLHGMYSATKEILIVQILATDPTKLDVRYSPFVKPKNVSTSHGSKNV 870
 QY 858 -----WVIERWGECSKCELGQWRRLVECRDINGQPASECAKEVPASTRPCADHPC 909
 DB 871 GSHTSQPQWVTGPWLACSTRCTDGTGWHTRTVQCGDNKRLAKGCPLSQRPSPAFKQCLLKKC 930
 RESULT 14
 US-09-634-286A-15
 ; Sequence 15, Application US/09634286A
 ; Patent No. 6521436
 ; GENERAL INFORMATION:
 ; APPLICANT: Bristol-Myers Squibb Company
 ; TITLE OF INVENTION: AGGRECAN DEGRADING METALLO PROTEASES
 ; FILE REFERENCE: DM6909A
 ; CURRENT APPLICATION NUMBER: US/09/634, 286A
 ; CURRENT FILING DATE: 2000-08-09
 ; NUMBER OF SEQ ID NOS: 21
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 15
 ; LENGTH: 930
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-634-286A-15
 * Query Match 36.4%; Score 1924; DB 4; Length 930;
 Best Local Similarity 40.7%; Pred. No. 2e-159;
 Matches 391; Conservative 125; Mismatches 314; Indels 130; Gaps 17;
 QY 25 PGRSFGPVPPTLLAAALLAVSDALGRPSEDEELVPELVRPGHG---TTLRLRHAF 81
 DB 26 PAQDKAGPPT-----AAAAQPRRQGEVEQERAEP-----PGHPPLAQRRSKGL 73
 QY 82 DQQLD-----LDVPPDSSFLAPGFTLQNVGRKSGSDTLPETDL 120
 DB 74 VQNIQLYSGGKGVLYVAGRRFLDLDERDGSVGIAGFV-----PAGGTSAPMRHR 127
 QY 121 AHCFVSGTVNGDPSSNAALSCEGVRGAFYLLGEAYFIQPL---PAASER----- 167
 DB 128 SHCFYRGTVDSAPRSLAVFDLCGGLDGFFAVKHARYTLKPLLRGPMAEERKGRVYDGS 187
 QY 168 -----LATAAPEKPPAPLQPHLLRRNQDVGCTGVVDDEP---RPTG 209
 DB 188 RILHYVTREGFSFEALPPRASCETPASTPEAH-----EHAPAHNSPG 230
 QY 210 KA-----ETBEDEGTEGEDEGPQ--WSPQDPALQGVQPTGTGTSIRKRFVSSHRYETM 263
 DB 231 RAALASQLLDQALSAGSGGFTWNR-----RRRSISRARQVELL 272
 QY 264 LVADQSMAPFHSGLKHYLLTFSVAARLYKHSIRNSVLSLVVVLVHDEQKPEVTS 323
 DB 273 LVADASMARLYGRGLQHYLLTASTANRLYSHASIHRIALVAVVVLGDKDKSLEYSK 332
 QY 324 NAALTILRNFCNWKQHNPPSDRAEYDTAILFTRODLGSGTCTDGLGWADYGVTCDDPR 383
 DB 333 NAATTILKFNCKWQHNLQDDEEYDAAIILFTREDLCGHHSCTDGLGWADYGVTCSPR 392
 QY 384 SCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACASLNGVNDQSHMMSMLNLDHSPW 443

DB 393 SCVIEDDGLHAAFTVAHEIIGHLGLSHDSDSKFCEETFGSTEDKRLMSSILTSIDASKPW 452
 QY 444 SPSCGYMITTSLDNGHGBCLMDKPNPTQLPGDLPGTSYDANROCOFTFGSDSKHCPDAA 503
 DB 453 SKCTSATITFELDDHGNCLLDLPRKQLLGPBEELPGQTYDATQCCNLTFGPEYSVCP-GM 511
 QY 504 STCSTLWCTGTSGGVLVCGQTKHFPWADGTSCEGKWCINGKCVNNHHRKHFDTPPHGWSG 563
 DB 512 DVCARLWCAVVRQGMVCLTKKLPAVEGTPCGKGRICLQKGCVDKTKKKYTSSTSHGNWG 571
 QY 564 MWGPWGDSSRRCCGGVQYTMRECDNPVKNCGKCEGKRVRYRSCNLEDCPDNNGKTPRE 623
 DB 572 SWGSGWQCSRCGGGVQPAYRHCHNNPAPRNNRGVCTGKRAIYRSCSLMPCPP-NGKSPRH 630
 QY 624 EQCAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLICQAKGIGYFFVLQPKVVDGTPC 683
 DB 631 ECEAKNGYQSDAKGVKTEFVWPKYAGVLPADVCKLTCRAKGTGYVVFSPKVTDGTEC 690
 QY 684 SPDSTSVCOGCVKAGCDRIIDSKKPKFCGCGNGSTCKKISGVSATSAKPGVHDIIT 743
 DB 691 RPYNSVVRGKCVRTGCGIIGSKLQYDKGCGGDNSSCTKIIVTFNKKSKGYTDVVR 750
 QY 744 IPTGATNIEVKORNGSRNNGSFLAIKAADGTIYILNGDYTLSTLEQDLMYKGVVLYRSG 803
 DB 751 IPEGATHIKVRQKAKQDRTFTAYLALKKNGEYLINGKYMISTETIIDINGTVMNYS 810
 QY 804 SSAALERIR--SFSPKLEPLTIQVLTVGNALRPKIYTYFV-KKKESFNAIPTFSA--- 857
 DB 811 WSHRDDFLHGMYSATKEILIVQILATDPTKLDVRYSPFVKPKNVSTSHGSKNV 870
 QY 858 -----WVIERWGECSKCELGQWRRLVECRDINGQPASECAKEVPASTRPCADHPC 909
 DB 871 GSHTSQPQWVTGPWLACSTRCTDGTGWHTRTVQCGDNKRLAKGCPLSQRPSPAFKQCLLKKC 930
 RESULT 15
 US-10-247-685-15
 ; Sequence 15, Application US/10247685
 ; Patent No. 6753176
 ; GENERAL INFORMATION:
 ; APPLICANT: Bristol-Myers Squibb Company
 ; TITLE OF INVENTION: AGGRECAN DEGRADING METALLO PROTEASES
 ; FILE REFERENCE: DM6909D
 ; CURRENT APPLICATION NUMBER: US/10/247,685
 ; CURRENT FILING DATE: 2002-09-19
 ; NUMBER OF SEQ ID NOS: 21
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 15
 ; LENGTH: 930
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-247-685-15
 Query Match 36.4%; Score 1924; DB 4; Length 930;
 Best Local Similarity 40.7%; Pred. No. 2e-159;
 Matches 391; Conservative 125; Mismatches 314; Indels 130; Gaps 17;
 QY 25 PGRSFGPVPPTLLAAALLAVSDALGRPSEDEELVPELVRPGHG---TTLRLRHAF 81
 DB 26 PAQDKAGPPT-----AAAAQPRRQGEVEQERAEP-----PGHPPLAQRRSKGL 73
 QY 82 DQQLD-----LDVPPDSSFLAPGFTLQNVGRKSGSDTLPETDL 120
 DB 74 VQNIQLYSGGKGVLYVAGRRFLDLDERDGSVGIAGFV-----PAGGTSAPMRHR 127
 QY 121 AHCFVSGTVNGDPSSNAALSCEGVRGAFYLLGEAYFIQPL---PAASER----- 167
 DB 128 SHCFYRGTVDSAPRSLAVFDLCGGLDGFFAVKHARYTLKPLLRGPMAEERKGRVYDGS 187
 QY 168 -----LATAAPEKPPAPLQPHLLRRNQDVGCTGVVDDEP---RPTG 209
 DB 188 RILHYVTREGFSFEALPPRASCETPASTPEAH-----EHAPAHNSPG 230

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2005, 10:12:49 ; Search time 167 Seconds
(without alignments)
2267.448 Million cell updates/sec

Title: US-09-989-687-126
Perfect score: 5287
Sequence: 1 MORAVPEGFGRKLGSDMGN.....CDPLKKPHFDICTMAECS 967

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1759131 seqs, 391586102 residues

Total number of hits satisfying chosen parameters: 1759131

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5287	100.0	967	11	US-09-989-687-126
2	5232	99.0	967	13	US-10-105-929-2
3	5232	99.0	967	14	US-10-115-286-2
4	5232	99.0	967	16	US-10-757-450-2
5	5234	98.8	967	16	US-10-755-889-134
6	5224	98.8	967	17	US-10-741-600-1603
7	5224	98.8	967	17	US-10-741-600-1604
8	5224	98.8	967	18	US-10-923-035-37
9	5221.5	98.8	968	10	US-09-373-658-125
10	5144	97.3	950	10	US-09-373-658-2
11	5144	97.3	950	11	US-09-989-687-2

12	5140	97.2	949	17	US-10-667-281-2	Sequence 2, Appli
13	5136	97.1	950	17	US-10-741-600-1605	Sequence 1605, Ap
14	5136	97.1	950	18	US-10-973-858-14	Sequence 14, Appl
15	5036	95.3	931	9	US-09-741-151-4	Sequence 4, Appli
16	4293.5	81.2	968	13	US-10-163-316-7	Sequence 7, Appli
17	4293.5	81.2	968	15	US-10-391-364-82	Sequence 82, Appl
18	4264.5	80.7	951	15	US-10-381-793-3	Sequence 3, Appli
19	4260.5	80.6	950	9	US-09-321-987B-4	Sequence 4, Appli
20	3922	74.2	727	9	US-09-445-023A-1	Sequence 1, Appli
21	3922	74.2	727	14	US-10-037-597-1	Sequence 1, Appli
22	3922	74.2	727	14	US-10-037-580-1	Sequence 12, Appl
23	3656	69.2	727	9	US-09-445-023A-12	Sequence 12, Appl
24	3656	69.2	727	14	US-10-037-597-12	Sequence 12, Appl
25	3656	69.2	727	14	US-10-037-580-12	Sequence 2, Appli
26	3303	62.5	608	9	US-09-803-589-2	Sequence 2, Appli
27	3303	62.5	608	16	US-10-718-332-2	Sequence 8, Appli
28	3297	62.4	608	13	US-09-803-589-8	Sequence 13, Appl
29	3297	62.4	608	13	US-10-105-929-13	Sequence 8, Appli
30	3297	62.4	608	16	US-10-718-332-8	Sequence 16, Appl
31	2794	52.8	551	9	US-09-802-582-16	Sequence 16, Appl
32	2794	52.8	551	13	US-10-105-929-16	Sequence 16, Appl
33	2794	52.8	551	14	US-10-365-227-16	Sequence 10, Appl
34	2635	49.8	518	9	US-09-803-589-10	Sequence 10, Appl
35	2635	49.8	518	16	US-10-718-332-10	Sequence 10, Appl
36	2495.5	47.2	924	15	US-10-033-463-28	Sequence 28, Appl
37	2478	46.9	978	15	US-10-275-107-59	Sequence 59, Appl
38	2477.5	46.9	950	9	US-09-741-151-2	Sequence 2, Appli
39	2477.5	46.9	950	9	US-09-965-631-4	Sequence 4, Appli
40	2477.5	46.9	950	16	US-10-391-364-77	Sequence 77, Appl
41	2477.5	46.9	950	17	US-10-753-267-56	Sequence 56, Appl
42	2477.5	46.9	950	17	US-10-961-020-4	Sequence 4, Appli
43	2471.5	46.7	950	16	US-10-763-210-1	Sequence 1, Appli
44	2467	46.7	577	15	US-10-425-114-39241	Sequence 39241, A
45	2455.5	46.4	952	15	US-10-311-035-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-09-989-687-126
; Sequence 126, Application US/09989687
; Publication No. US20040002449A1
; GENERAL INFORMATION:
; APPLICANT: Hastings, Gregg A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Meth1 and Meth2 Polynucleotides and Polypeptides
; FILE REFERENCE: 1488.107000D
; CURRENT APPLICATION NUMBER: US/09/989,687
; CURRENT FILING DATE: 2001-11-21
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 126
; LENGTH: 967
; TYPE: PRT
; ORGANISM: ITGL-TSP
US-09-989-687-126

Query Match 100.0%; Score 5287; DB 11; Length 967;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 967; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MORAVPEGFGRKLGSDMGNAPGSRSGFVPPTLLLAALLAVSDALGRPSDEEL	60
Db	1	MORAVPEGFGRKLGSDMGNAPGSRSGFVPPTLLLAALLAVSDALGRPSDEEL	60
Qy	61	VPELVRVPGHGTTRLRHAFDQQLDLDVPPDSSFLAGFTLQNVGRKSGSDTLPETDL	120
Db	61	VPELVRVPGHGTTRLRHAFDQQLDLDVPPDSSFLAGFTLQNVGRKSGSDTLPETDL	120
Qy	121	AHCFYSGTVNGDPSSAAALSLCEGVGRGAFYLLGEAFYIQLPAAERLATAPGKPPAP	180
Db	121	AHCFYSGTVNGDPSSAAALSLCEGVGRGAFYLLGEAFYIQLPAAERLATAPGKPPAP	180

```
QY 181 LQFHLLRRNRQDVGCTCGVVDDEPRPTGKARTEDDEDETEDEGPGQSPDPAALQGVG 240
Db 181 LQFHLLRRNRQDVGCTCGVVDDEPRPTGKARTEDDEDETEDEGPGQSPDPAALQGVG 240
QY 241 QPTGTGSIKKRKFVSSHRYVETMLVADQSMAEFHGSLGKHYLLTLFVSAARLYKHPISRN 300
Db 241 QPTGTGSIKKRKFVSSHRYVETMLVADQSMAEFHGSLGKHYLLTLFVSAARLYKHPISRN 300
QY 301 SVSLVVVKILVTHDQKQPEVTSNAALTLRPNFCWQKQHNPPSDRDAEHYDTAILFTTQD 360
Db 301 SVSLVVVKILVTHDQKQPEVTSNAALTLRPNFCWQKQHNPPSDRDAEHYDTAILFTTQD 360
QY 361 LCGSQTCDTLGMADVGTVCDSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACSL 420
Db 361 LCGSQTCDTLGMADVGTVCDSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACSL 420
QY 421 NGVNQDSHMMASMLNLDHSPWSPCSGYMITSLFDNGHGCECLMDKPNPQLPGDLPGT 480
Db 421 NGVNQDSHMMASMLNLDHSPWSPCSGYMITSLFDNGHGCECLMDKPNPQLPGDLPGT 480
QY 481 SYDANROCOFTFGEDSKHCIPDAASTCSTLWCTGTSGGVLVCQTKHFPWADGTSCBEGKWC 540
Db 481 SYDANROCOFTFGEDSKHCIPDAASTCSTLWCTGTSGGVLVCQTKHFPWADGTSCBEGKWC 540
QY 541 INKCVNKNHRKHFDTPFHGSGWGMWPGWDCSRTCCTGGGVQYTMRECDNPVPKNGKCYCEG 600
Db 541 INKCVNKNHRKHFDTPFHGSGWGMWPGWDCSRTCCTGGGVQYTMRECDNPVPKNGKCYCEG 600
QY 601 KRVRYSNLEDCPDNNGKTFREEQCEAHNEFASKASFGSGPAVEWIPKYAGVSPKDRCKL 660
Db 601 KRVRYSNLEDCPDNNGKTFREEQCEAHNEFASKASFGSGPAVEWIPKYAGVSPKDRCKL 660
QY 661 ICQAKGIGYFFVLQPKVVDGTPCSPDSTSVCVQGCVKAGCDRIIDSKKKFDKCGVCGGN 720
Db 661 ICQAKGIGYFFVLQPKVVDGTPCSPDSTSVCVQGCVKAGCDRIIDSKKKFDKCGVCGGN 720
QY 721 GSTCKKISGSVTSKAPGYHDIITPTGATNIEVKORNGSRNNGSFIAIKAADGTYILN 780
Db 721 GSTCKKISGSVTSKAPGYHDIITPTGATNIEVKORNGSRNNGSFIAIKAADGTYILN 780
QY 781 GDYTLSTLEQDIIMYKGVVLYRSGSSAALERIRSFPLKEPLTIQVLTGVGNALRPKIKYTY 840
Db 781 GDYTLSTLEQDIIMYKGVVLYRSGSSAALERIRSFPLKEPLTIQVLTGVGNALRPKIKYTY 840
QY 841 FVKKKKESFNAIPTFSAWVIEEWGECSSKSELGWQRRLVECRDINGQPASECAKEVKPAS 900
Db 841 FVKKKKESFNAIPTFSAWVIEEWGECSSKSELGWQRRLVECRDINGQPASECAKEVKPAS 900
```

RESULT 2

```
US-10-105-929-2
; Sequence 2, Application US/10105929
; Publication No. US20020137142A1
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
; FILE REFERENCE: 09404/041001
; CURRENT APPLICATION NUMBER: US/10/105,929
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/130,491
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-07
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/058,108
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-05
```

```
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/054,961
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 967
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-105-929-2
```

```
Query Match 99.0%; Score 5232; DB 13; Length 967;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 957; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 MORAVPFGFGRKLGSDMGNAERAPGSRSGFVPTLLALLAAALLAVSDALGRPSEEBEL 60
Db 1 MORAVPFGFGRKLGSDMGNAERAPGSRSGFVPTLLALLAAALLAVSDALGRPSEEBEL 60
QY 61 VVPELERVPVGHGTTRLRLHAFDQQLDLDVPPDSSFLAPGFTLQNVGRKSGSDTLPETDL 120
Db 61 VVPELERVPVGHGTTRLRLHAFDQQLDLDVPPDSSFLAPGFTLQNVGRKSGSETPLPETDL 120
QY 121 AHCFSYGTVNGDPSSAAALSCEGVRGAFYLLGBAYFIQPLPAASERLATAAPGEKPPAP 180
Db 121 AHCFSYGTVNGDPSSAAALSCEGVRGAFYLLGBAYFIQPLPAASERLATAAPGEKPPAP 180
QY 181 LQFHLLRRNRQDVGCTCGVVDDEPRPTGKARTEDDEDETEDEGPGQSPDPAALQGVG 240
Db 181 LQFHLLRRNRQDVGCTCGVVDDEPRPTGKARTEDDEDETEDEGPGQSPDPAALQGVG 240
QY 241 QPTGTGSIKKRKFVSSHRYVETMLVADQSMAEFHGSLGKHYLLTLFVSAARLYKHPISRN 300
Db 241 QPTGTGSIKKRKFVSSHRYVETMLVADQSMAEFHGSLGKHYLLTLFVSAARLYKHPISRN 300
QY 301 SVSLVVVKILVTHDQKQPEVTSNAALTLRPNFCWQKQHNPPSDRDAEHYDTAILFTTQD 360
Db 301 SVSLVVVKILVTHDQKQPEVTSNAALTLRPNFCWQKQHNPPSDRDAEHYDTAILFTTQD 360
QY 361 LCGSQTCDTLGMADVGTVCDSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACSL 420
Db 361 LCGSQTCDTLGMADVGTVCDSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACSL 420
QY 421 NGVNQDSHMMASMLNLDHSPWSPCSGYMITSLFDNGHGCECLMDKPNPQLPGDLPGT 480
Db 421 NGVNQDSHMMASMLNLDHSPWSPCSAYMTSLFDNGHGCECLMDKPNPQLPGDLPGT 480
QY 481 SYDANROCOFTFGEDSKHCIPDAASTCSTLWCTGTSGGVLVCQTKHFPWADGTSCBEGKWC 540
Db 481 SYDANROCOFTFGEDSKHCIPDAASTCSTLWCTGTSGGVLVCQTKHFPWADGTSCBEGKWC 540
QY 541 INKCVNKNHRKHFDTPFHGSGWGMWPGWDCSRTCCTGGGVQYTMRECDNPVPKNGKCYCEG 600
Db 541 INKCVNKNTRDKHFDTPFHGSGWGMWPGWDCSRTCCTGGGVQYTMRECDNPVPKNGKCYCEG 600
QY 601 KRVRYSNLEDCPDNNGKTFREEQCEAHNEFASKASFGSGPAVEWIPKYAGVSPKDRCKL 660
Db 601 KRVRYSNLEDCPDNNGKTFREEQCEAHNEFASKASFGSGPAVEWIPKYAGVSPKDRCKL 660
QY 661 ICQAKGIGYFFVLQPKVVDGTPCSPDSTSVCVQGCVKAGCDRIIDSKKKFDKCGVCGGN 720
Db 661 ICQAKGIGYFFVLQPKVVDGTPCSPDSTSVCVQGCVKAGCDRIIDSKKKFDKCGVCGGN 720
QY 721 GSTCKKISGSVTSKAPGYHDIITPTGATNIEVKORNGSRNNGSFIAIKAADGTYILN 780
Db 721 GSTCKKISGSVTSKAPGYHDIITPTGATNIEVKORNGSRNNGSFIAIKAADGTYILN 780
QY 781 GDYTLSTLEQDIIMYKGVVLYRSGSSAALERIRSFPLKEPLTIQVLTGVGNALRPKIKYTY 840
Db 781 GDYTLSTLEQDIIMYKGVVLYRSGSSAALERIRSFPLKEPLTIQVLTGVGNALRPKIKYTY 840
QY 841 FVKKKKESFNAIPTFSAWVIEEWGECSSKSELGWQRRLVECRDINGQPASECAKEVKPAS 900
Db 841 FVKKKKESFNAIPTFSAWVIEEWGECSSKSELGWQRRLVECRDINGQPASECAKEVKPAS 900
```

QY 901 TRPCADHPCPQWLGWSSCKTCGKYKTSKLSHSDGGVLSHSDCDPLKKPKHFIDF 960
Db 901 TRPCADHPCPQWLGWSSCKTCGKYKTSKLSHSDGGVLSHSDCDPLKKPKHFIDF 960
QY 961 CTMAECS 967
Db 961 CTMAECS 967

RESULT 3
US-10-115-286-2
; Sequence 2, Application US/10115286
; Publication No. US20030166065A1
; GENERAL INFORMATION:
; APPLICANT: Jonak, Zdenka
; Trulli, Stephen
; Fornwald, James
; Terrett, Jonathan
; Hastings, Gregg
; TITLE OF INVENTION: No. US20030166065A1e1 Integrin Ligand ITGL-TSP
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ratner & Prestia
; STREET: Box 980
; CITY: Valley Forge
; STATE: PA
; COUNTRY: USA
; ZIP: 19482

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/115,286
FILING DATE: 04-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/845,496
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Prestia, Paul F.
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 967 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-115-286-2

Query Match 99.0%; Score 5232; DB 14; Length 967;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 957; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 MORAVPEGFRKRLGSDMGNAERAPGSRFGVPVPTLLLLAALLAVSDALGRPSSEDEEL 60
Db 1 MORAVPEGFRKRLGSDMGNAERAPGSRFGVPVPTLLLLAALLAVSDALGRPSSEDEEL 60
QY 61 VYPELVRPGHGTTRLRHAFQQLDLVPPDSSFLAPGFTLQNVGRKSGSDTLPETDL 120
Db 61 VYPELERAPGHGTTRLRHAFQQLDLRLPDSSFLAPGFTLQNVGRKSGSETPLPETDL 120
QY 121 AHCFYSGTVNGDPSAAALSLCEGVRGAFYLLGEAYFIQPLPAASERLATAPGKPPAP 180

Db 121 AHCFYSGTVNGDPSAAALSLCEGVRGAFYLLGEAYFIQPLPAASERLATAPGKPPAP 180
QY 181 LQFHLLRNRQDVGCTCGVVDDEPRPTGKAETDEDEGEDEGGPQWSPQDPALQGVG 240
Db 181 LQFHLLRNRQDVGCTCGVVDDEPRPTGKAETDEDEGEDEGGPQWSPQDPALQGVG 240
QY 241 OPTGTGSIIRKKRFVSSSHRYVETMLVADQSMAEFGSLGKHYLLTLFSAARLYKHPISRN 300
Db 241 OPTGTGSIIRKKRFVSSSHRYVETMLVADQSMAEFGSLGKHYLLTLFSAARLYKHPISRN 300
QY 301 SVSLVVVKILVIHDEOKGPEVTSNAALTLRNFCWKQKOHNPSPDRDASHYDTAILFTTQD 360
Db 301 SVSLVVVKILVIHDEOKGPEVTSNAALTLRNFCWKQKOHNPSPDRDASHYDTAILFTTQD 360
QY 361 LCGSQTCDTLGMADVGTVCDFPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACSL 420
Db 361 LCGSQTCDTLGMADVGTVCDFPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACSL 420
QY 421 NGVNQDSHMMASMLSLNLDHSPQWSPFCSGYMTSFLDNHGHGCLMDKPNPIQLPGLDLPGT 480
Db 421 NGVNQDSHMMASMLSLNLDHSPQWSPFCSGYMTSFLDNHGHGCLMDKPNPIQLPGLDLPGT 480
QY 481 SYDANRQCFQTFGSDSKHCPDAASTCTSLWCTGTSGGVLVCOTKHPWADGTSCGEGKWC 540
Db 481 SYDANRQCFQTFGSDSKHCPDAASTCTSLWCTGTSGGVLVCOTKHPWADGTSCGEGKWC 540
QY 541 INKCVNKNHRKHFDTDPFHGSGWGMWPGDCSRTCGGQVQYTMRECDNPFVKNKGKCYEG 600
Db 541 INKCVNKNHRKHFDTDPFHGSGWGMWPGDCSRTCGGQVQYTMRECDNPFVKNKGKCYEG 600
QY 601 KRVYRSCNLEDCPDNNGKTFREQCCEAHNPFSAKSGSGPAVEWI PKYAGVSPKDRCKL 660
Db 601 KRVYRSCNLEDCPDNNGKTFREQCCEAHNPFSAKSGSGPAVEWI PKYAGVSPKDRCKL 660
QY 661 ICQAKGIGYFFVLQPKVVDGTPCSPDSTSCVQCGVKAGCDRIIDSKKFKDKGVCQGN 720
Db 661 ICQAKGIGYFFVLQPKVVDGTPCSPDSTSCVQCGVKAGCDRIIDSKKFKDKGVCQGN 720
QY 721 GSTCKKISGSVTSAPKGYVDIITPTGATNIEVQRNQRSGRNNGSFLAIKAADGTYILN 780
Db 721 GSTCKKISGSVTSAPKGYVDIITPTGATNIEVQRNQRSGRNNGSFLAIKAADGTYILN 780
QY 781 GDYTLSTLEQDQIMYKGVVLYSGSSAALERIRSPKLEPTIQVLTGVGNALRPKIKYTY 840
Db 781 GDYTLSTLEQDQIMYKGVVLYSGSSAALERIRSPKLEPTIQVLTGVGNALRPKIKYTY 840
QY 841 FVKKKKESFNAIPTFSAWVIEEWGECSSKCELGWQRRLVECRDINGOPASCAKEVPKAS 900
Db 841 FVKKKKESFNAIPTFSAWVIEEWGECSSKCELGWQRRLVECRDINGOPASCAKEVPKAS 900
QY 901 TRPCADHPCPQWLGWSSCKTCGKYKTSKLSHSDGGVLSHSDCDPLKKPKHFIDF 960
Db 901 TRPCADHPCPQWLGWSSCKTCGKYKTSKLSHSDGGVLSHSDCDPLKKPKHFIDF 960
QY 961 CTMAECS 967
Db 961 CTMAECS 967

RESULT 4
US-10-757-450-2
; Sequence 2, Application US/10757450
; Publication No. US20040175794A1
; GENERAL INFORMATION:
; APPLICANT: Jonak, Zdenka
; Trulli, Stephen
; Fornwald, James
; Terrett, Jonathan
; Hastings, Gregg
; TITLE OF INVENTION: Novel Integrin Ligand ITGL-TSP
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ratner & Prestia


```

; STREET: Box 980
; CITY: Valley Forge
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Fast-SEQ for Windows Version 2.0
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/757,450
; FILING DATE: 15-Jan-2004
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/10/115,286
; FILING DATE: 04-Apr-2002
; APPLICATION NUMBER: 08/845,496
; FILING DATE: <Unknown>
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Prestia, Paul F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
;
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 967 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
;
; US-10-757-450-2

```

Query Match	99.0%;	Score 5232;	DB 16;	Length 967;
Best Local Similarity	99.0%;	Pred. No. 0;		
Matches 957;	Conservative 4;	Mismatches 6;	Indels 0;	Gaps 0;
Qy	1	MQRVAPGFGRRKLGSDMGNAERAPGSRSGFVPFTLLLLAAALLAVSDALGRPSEDEEL	60	
Db	1	MQRVAPGFGRRKLGSDMGNAERAPGSRSGFVPFTLLLLAAALLAVSDALGRPSEDEEL	60	
Qy	61	VPELERVPOHGHTTRLRLHAFDQQLDLDDPPDSSFLAPGFTLQVGRKSGSDTLPETDL	120	
Db	61	VPELERAPGHGHTTRLRLHAFDQQLDLLELRDSSFLAPGFTLQVGRKSGSETLPETDL	120	
Qy	121	AHCFYSGTVNGDPSSAAALSCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGEKEPPAP	180	
Db	121	AHCFYSGTVNGDPSSAAALSCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGEKEPPAP	180	
Qy	181	LQPHLLARNRQGDVGTCGVVDDDEPRPTGKAETDEDEGEDEGEQWSPQDPALQGVG	240	
Db	181	LQPHLLARNRQGDVGTCGVVDDDEPRPTGKAETDEDEGEDEGEQWSPQDPALQGVG	240	
Qy	241	OPTGTGSGIRKKRFVSSHRYVETMLVADQSMAEFHGSLKHVLLTLFVSAARLYKHPISRN	300	
Db	241	OPTGTGSGIRKKRFVSSHRYVETMLVADQSMAEFHGSLKHVLLTLFVSAARLYKHPISRN	300	
Qy	301	SVSLVWVKILVITHDEQKGPEVTSNAAULTLRNFCNWQKOHNPSPDRDAEHYDVTALFLTRQD	360	
Db	301	SVSLVWVKILVITHDEQKGPEVTSNAAULTLRNFCNWQKOHNPSPDRDAEHYDVTALFLTRQD	360	
Qy	361	LCGSGTCDDTLGMDADVGVCPDPSRSCSVIEDDGLQAAFTTAHELGHVFMNPHDDAKQCASL	420	
Db	361	LCGSGTCDDTLGMDADVGVCPDPSRSCSVIEDDGLQAAFTTAHELGHVFMNPHDDAKQCASL	420	
Qy	421	NGVNQDSHMMASMLSNLDHSPQSPSCSGYMITSFLDNGHGCECLMDKQNPQLPGDLPGT	480	
Db	421	NGVNQDSHMMASMLSNLDHSPQSPSCSGYMITSFLDNGHGCECLMDKQNPQLPGDLPGT	480	

Qy	481	SYDANROCOFTFGEDSKHCPDAASTCSLTWCTGTSGGVLVQCTKHPWADGTSCEGKWC	540
Db	481	SYDANROCOFTFGEDSKHCPDAASTCSLTWCTGTSGGVLVQCTKHPWADGTSCEGKWC	540
Qy	541	INGKCVNKNHRKHFDTPFHGSGWGMWPGWDCSRTCGGGVQVYTMRECDNPVPKNGGKYCBG	600
Db	541	INGKCVNKNTRKHFDTPFHGSGWGMWPGWDCSRTCGGGVQVYTMRECDNPVPKNGGKYCBG	600
Qy	601	KRVRYRSCNLEDCPDNRNGKTFREOCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKL	660
Db	601	KRVRYRSCNLEDCPDNRNGKTFREOCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKL	660
Qy	661	ICQAKGIGYFVLPQKVVDGTPCSPDSTSVQVQGCYKAGCDRIIDSKKKFDKCGVCGN	720
Db	661	ICQAKGIGYFVLPQKVVDGTPCSPDSTSVQVQGCYKAGCDRIIDSKKKFDKCGVCGN	720
Qy	721	GSTCKKISGVSATSAKPGYHDIITPTGATNIEVKQNRQSRNNGSFLLAIKAADGTIYL	780
Db	721	GSTCKKISGVSATSAKPGYHDIITPTGATNIEVKQNRQSRNNGSFLLAIKAADGTIYL	780
Qy	781	GDYTLSTLEQIMYKGVVLYRSGSSAALERIRSFSPKLEPTIQVLTVGNALRPKIKTY	840
Db	781	GDYTLSTLEQIMYKGVVLYRSGSSAALERIRSFSPKLEPTIQVLTVGNALRPKIKTY	840
Qy	841	FVKKKESFNAIPTFSAWLIEEWCECKSCBLGHWRRLVECRDINGQPASECAKEVKPAS	900
Db	841	FVKKKESFNAIPTFSAWLIEEWCECKSCBLGHWRRLVECRDINGQPASECAKEVKPAS	900
Qy	901	TRPCADHPCQWQLGEWSSCKTCGKYKTSLSKLSHDGVLSHSDCDPLKKPKHFIDF	960
Db	901	TRPCADHPCQWQLGEWSSCKTCGKYKTSLSKLSHDGVLSHSDCDPLKKPKHFIDF	960
Qy	961	CTWAECS 967	
Db	961	CTWAECS 967	

RESULT 5

US-10-755-889-134

; Sequence 134, Application US/10755889

; Publication No. US20040171823A1

; GENERAL INFORMATION:

; APPLICANT: Bristol-Myers Squibb Company

; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-k-B

; TITLE OF INVENTION: PATHWAY

; FILE REFERENCE: D0284 NP

; CURRENT APPLICATION NUMBER: US/10/755,889

; CURRENT FILING DATE: 2004-01-13

; PRIOR APPLICATION NUMBER: U.S. 60/440,068

; PRIOR FILING DATE: 2003-01-14

; PRIOR APPLICATION NUMBER: U.S. 60/469,757

; PRIOR FILING DATE: 2003-05-12

; NUMBER OF SEQ ID NOS: 823

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 134

; LENGTH: 967

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-755-889-134

Query Match		98.8%	Score 5224;	DB 16;	Length 967;
Best Local Similarity		98.8%	Pred. No. 0;		
Matches	955;	Conservative	5;	Mismatches	7;
				Indels	0;
				Gaps	0;

Qy	1	MQRAVPEGFGRKRLGSDMGNAERAPGSRSGFVPPTLLLLAALLAVSDALGRPSEDEEL	60
Db	1	MQRAVPEGFGRKRLGSDMGNAERAPGSRSGFVPPTLLLLAALLAVSDALGRPSEDEEL	60
Qy	61	VPELELRVPGHGTTRLRHLHAFDQOLDLVDPPDSSFLAPGFTLQNVGRKSGSDTLPETDL	120
Db	61	VPELELRAPGHGTTRLRHLHAFDQOLDLVDPPDSSFLAPGFTLQNVGRKSGSETLPETDL	120
Qy	121	AHCFYSGTVNGDPSFSSAALSLCEGVRAFYLGLGAYFIQPLPAASERLATAPGKPPAP	180

```
Db 121 AHCFYSGTVNGDPSSAAALSLCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGKPPAP 180
Qy 181 LQFHLLRRNRQDVGTCGVVDDEPRPTGKATEDEDETEGEGPQSPDPAALQGVG 240
Db 181 LQFHLLRRNRQDVGTCGVVDDEPRPTGKATEDEDETEGEGPQSPDPAALQGVG 240
Qy 241 QPTGTGSRKKRKFVSSHRYVETMLVADQSMABFHGSLGHYLLTLFSAARLYKHPSIRN 300
Db 241 QPTGTGSRKKRKFVSSHRYVETMLVADQSMABFHGSLGHYLLTLFSAARLYKHPSIRN 300
Qy 301 SVSLVVVKILVTHDQKGPEVTSNAALTILRNFCNWKQHNPPSDRDAEHYDTAILFTQD 360
Db 301 SVSLVVVKILVTHDQKGPEVTSNAALTILRNFCNWKQHNPPSDRDAEHYDTAILFTQD 360
Qy 361 LCGSOTCDTLGNADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPNHDDAKQACSL 420
Db 361 LCGSOTCDTLGNADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPNHDDAKQACSL 420
Qy 421 NGVNQDSHMMASMLNLDHSPWSPCSGYMITSLDNGHGECIMDKPQNPILQPGDLPGT 480
Db 421 NGVNQDSHMMASMLNLDHSPWSPCSGYMITSLDNGHGECIMDKPQNPILQPGDLPGT 480
Qy 481 SYDANRQOQFTFGEDSKHCPDAASTCTLWCTGTSGVLVVCOTKHPPWADGTSCGEGKWC 540
Db 481 SYDANRQOQFTFGEDSKHCPDAASTCTLWCTGTSGVLVVCOTKHPPWADGTSCGEGKWC 540
Qy 541 INKCVNKNHKKHFDTPFHGSGMWGMPGDCSRTCGGVQVYTMRECDNPVPKNGSKYCEG 600
Db 541 INKCVNKNHKKHFDTPFHGSGMWGMPGDCSRTCGGVQVYTMRECDNPVPKNGSKYCEG 600
Qy 601 KRVRYSNCLDCPDNNGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKL 660
Db 601 KRVRYSNCLDCPDNNGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKL 660
Qy 661 ICQAKGIGYFFVLQPKVVDGTPCSDPSTSVVQGCVKAGCDRIIDSKKKPKCGVCGN 720
Db 661 ICQAKGIGYFFVLQPKVVDGTPCSDPSTSVVQGCVKAGCDRIIDSKKKPKCGVCGN 720
Qy 721 GSTCKKISGSVTSAPGYHDIITPTGATNIEVKQNRQGRNNGSFLAIKAADGTIYN 780
Db 721 GSTCKKISGSVTSAPGYHDIITPTGATNIEVKQNRQGRNNGSFLAIKAADGTIYN 780
Qy 781 GDTLSTLEQDIMYKGVVLYRSGSSAALERIRSFSPKLEPLTIQVLTGNALRPKIKYTY 840
Db 781 GDTLSTLEQDIMYKGVVLYRSGSSAALERIRSFSPKLEPLTIQVLTGNALRPKIKYTY 840
Qy 841 FVKKKESFNAIPTFSAWVIEBWGECSSKSELGWQRRLVECRDINGQPASECAKEVKPAS 900
Db 841 FVKKKESFNAIPTFSAWVIEBWGECSSKSELGWQRRLVECRDINGQPASECAKEVKPAS 900
Qy 901 TRPCADHPCPQWLGEWSSCKTCGKYKTSKLCLSHDGGVLSHSDCDPLKKPKHFD 960
Db 901 TRPCADHPCPQWLGEWSSCKTCGKYKTSKLCLSHDGGVLSHSDCDPLKKPKHFD 960
```

RESULT 6

```
US-10-741-600-1603
; Sequence 1603, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 1603
; LENGTH: 967
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-600-1603

Query Match 98.8%; Score 5224; DB 17; Length 967;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 956; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 1 MORAPEGFGRRKLGSDMGNAERAPGSRSGFVPVPTLLLLAAALLAVSDALGRPSSEDEL 60
Db 1 MORAPEGFGRRKLGSDMGNAERAPGSRSGFVPVPTLLLLAAALLAVSDALGRPSSEDEL 60
Qy 61 VPELERVPGHGTTTLRLHAFDQQLDLDVDPDSFPAFGFTLQNVGRKSGSDTLPETDL 120
Db 61 VPELERAPGHGTTTLRLHAFDQQLDLELRPDSSFLAPGFTLQNVGRKSGSETLPETDL 120
Qy 121 AHCFYSGTVNGDPSSAAALSLCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGKPPAP 180
Db 121 AHCFYSGTVNGDPSSAAALSLCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGKPPAP 180
Qy 181 LQFHLLRRNRQDVGTCGVVDDEPRPTGKATEDEDETEGEGPQSPDPAALQGVG 240
Db 181 LQFHLLRRNRQDVGTCGVVDDEPRPTGKATEDEDETEGEGPQSPDPAALQGVG 240
Qy 241 QPTGTGSRKKRKFVSSHRYVETMLVADQSMABFHGSLGHYLLTLFSAARLYKHPSIRN 300
Db 241 QPTGTGSRKKRKFVSSHRYVETMLVADQSMABFHGSLGHYLLTLFSAARLYKHPSIRN 300
Qy 301 SVSLVVVKILVTHDQKGPEVTSNAALTILRNFCNWKQHNPPSDRDAEHYDTAILFTQD 360
Db 301 SVSLVVVKILVTHDQKGPEVTSNAALTILRNFCNWKQHNPPSDRDAEHYDTAILFTQD 360
Qy 361 LCGSOTCDTLGNADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPNHDDAKQACSL 420
Db 361 LCGSOTCDTLGNADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPNHDDAKQACSL 420
Qy 421 NGVNQDSHMMASMLNLDHSPWSPCSGYMITSLDNGHGECIMDKPQNPILQPGDLPGT 480
Db 421 NGVNQDSHMMASMLNLDHSPWSPCSGYMITSLDNGHGECIMDKPQNPILQPGDLPGT 480
Qy 481 SYDANRQOQFTFGEDSKHCPDAASTCTLWCTGTSGVLVVCOTKHPPWADGTSCGEGKWC 540
Db 481 SYDANRQOQFTFGEDSKHCPDAASTCTLWCTGTSGVLVVCOTKHPPWADGTSCGEGKWC 540
Qy 541 INKCVNKNHKKHFDTPFHGSGMWGMPGDCSRTCGGVQVYTMRECDNPVPKNGSKYCEG 600
Db 541 INKCVNKNHKKHFDTPFHGSGMWGMPGDCSRTCGGVQVYTMRECDNPVPKNGSKYCEG 600
Qy 601 KRVRYSNCLDCPDNNGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKL 660
Db 601 KRVRYSNCLDCPDNNGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKL 660
Qy 661 ICQAKGIGYFFVLQPKVVDGTPCSDPSTSVVQGCVKAGCDRIIDSKKKPKCGVCGN 720
Db 661 ICQAKGIGYFFVLQPKVVDGTPCSDPSTSVVQGCVKAGCDRIIDSKKKPKCGVCGN 720
Qy 721 GSTCKKISGSVTSAPGYHDIITPTGATNIEVKQNRQGRNNGSFLAIKAADGTIYN 780
Db 721 GSTCKKISGSVTSAPGYHDIITPTGATNIEVKQNRQGRNNGSFLAIKAADGTIYN 780
Qy 781 GDTLSTLEQDIMYKGVVLYRSGSSAALERIRSFSPKLEPLTIQVLTGNALRPKIKYTY 840
Db 781 GDTLSTLEQDIMYKGVVLYRSGSSAALERIRSFSPKLEPLTIQVLTGNALRPKIKYTY 840
Qy 841 FVKKKESFNAIPTFSAWVIEBWGECSSKSELGWQRRLVECRDINGQPASECAKEVKPAS 900
Db 841 FVKKKESFNAIPTFSAWVIEBWGECSSKSELGWQRRLVECRDINGQPASECAKEVKPAS 900
Qy 901 TRPCADHPCPQWLGEWSSCKTCGKYKTSKLCLSHDGGVLSHSDCDPLKKPKHFD 960
Db 901 TRPCADHPCPQWLGEWSSCKTCGKYKTSKLCLSHDGGVLSHSDCDPLKKPKHFD 960
```

```
Qy 961 CTMAECS 967
Db 961 CTMAECS 967

RESULT 7
US-10-741-600-1604
; Sequence 1604, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1604
; LENGTH: 967
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-600-1604

Query Match 98.8%; Score 5224; DB 17; Length 967;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 956; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 1 MORAVPEGFGRKRLGSDMGNAERAPGSRSGFVPPTLLLLAAALLAVSDALGRPSEDEEL 60
Db 1 MORAVPEGFGRKRLGSDMGNAERAPGSRSGFVPPTLLLLAAALLAVSDALGRPSEDEEL 60

Qy 61 VPELERVPCHGTTTLRLHAFDQDLDPDSSFLAPGFTLQNVGRKSGSDTLPETDL 120
Db 61 VPELERVPCHGTTTLRLHAFDQDLDPDSSFLAPGFTLQNVGRKSGSETLPETDL 120

Qy 121 AHCFYSGTVNGDPSSAAALSCEGVGAFYLLGGEAYFIQPLPAASERLATAAPEKPPAP 180
Db 121 AHCFYSGTVNGDPSSAAALSCEGVGAFYLLGGEAYFIQPLPAASERLATAAPEKPPAP 180

Qy 181 LQFHLLRRNQDVGTCGVVDDEPRPTGKAETDEDEGEGEPQWSPDPAALQGVG 240
Db 181 LQFHLLRRNQDVGTCGVVDDEPRPTGKAETDEDEGEGEPQWSPDPAALQGVG 240

Qy 241 QPTGTGSIKRRKRVSSHRYVETMLVADQSMAPFHGSLKHVLLTLFVVAARLYKHPISRN 300
Db 241 QPTGTGSIKRRKRVSSHRYVETMLVADQSMAPFHGSLKHVLLTLFVVAARLYKHPISRN 300

Qy 301 SVSLVVVKILVIHDEQKGPVTSNAALTFLRNFQWQKHNPPSDRDAEHYDTAILFTRQD 360
Db 301 SVSLVVVKILVIHDEQKGPVTSNAALTFLRNFQWQKHNPPSDRDAEHYDTAILFTRQD 360

Qy 361 LCGSQTCDTLGMADVGTVCDSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACSL 420
Db 361 LCGSQTCDTLGMADVGTVCDSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACSL 420

Qy 421 NGVNQDSHMAASMLNLDHSPFWSPCSGYMITSLDNGHGECGLMDKQNPQLPGLDLPGT 480
Db 421 NGVNQDSHMAASMLNLDHSPFWSPCSGYMITSLDNGHGECGLMDKQNPQLPGLDLPGT 480

Qy 481 SYDANRQCOFTFGEDSKHCPDAASTCTLWCTGTSGGVLCVQTKHPWADGTCGEGKWC 540
Db 481 SYDANRQCOFTFGEDSKHCPDAASTCTLWCTGTSGGVLCVQTKHPWADGTCGEGKWC 540

Qy 541 INKCVNKNHRKHFDPTPHGSGWMGMPGDCSRTC GGGVQVTMRECDNVPKNGKCYCEG 600
Db 541 INKCVNKNHRKHFDPTPHGSGWMGMPGDCSRTC GGGVQVTMRECDNVPKNGKCYCEG 600

Qy 601 KRVYRSCNLEDPCDNNKGTFRBQCEAHNEFSAKSCSGSPAVSEWIPKYAGVSPKDRCKL 660
Db 601 KRVYRSCNLEDPCDNNKGTFRBQCEAHNEFSAKSCSGSPAVSEWIPKYAGVSPKDRCKL 660
```

```
Qy 661 ICQAKGIGYFFVLQPKVVDGTPCSPDSTSVQCQCVCAGCDRIIDSKKFKDKGVCQGN 720
Db 661 ICQAKGIGYFFVLQPKVVDGTPCSPDSTSVQCQCVCAGCDRIIDSKKFKDKGVCQGN 720

Qy 721 GSTCKKISGSVTSAPKPGYHDIITITPTGATNIEVKORNQSRNNGSFLAIKAADGTIYN 780
Db 721 GSTCKKISGSVTSAPKPGYHDIITITPTGATNIEVKORNQSRNNGSFLAIKAADGTIYN 780

Qy 781 GDYTLSTLEQDQIMYKGVLYRYSGSSAALERIRSFSPLEKEPTIQVLTGVALRPKIKYTY 840
Db 781 GDYTLSTLEQDQIMYKGVLYRYSGSSAALERIRSFSPLEKEPTIQVLTGVALRPKIKYTY 840

Qy 841 FVKKKESFNAIPTFSAWTIEEMCEKSCSCELGHQRRRLVECRDINGOPASECAKEVPAS 900
Db 841 FVKKKESFNAIPTFSAWTIEEMCEKSCSCELGHQRRRLVECRDINGOPASECAKEVPAS 900

Qy 901 TRPCADHPCQWQLGEWSSCKTCGKGYKRSKLSKLSHSDGVLSHSDCLPKPKHFIDF 960
Db 901 TRPCADHPCQWQLGEWSSCKTCGKGYKRSKLSKLSHSDGVLSHSDCLPKPKHFIDF 960

Qy 961 CTMAECS 967
Db 961 CTMAECS 967

RESULT 8
US-10-923-035-37
; Sequence 37, Application US/10923035
; Publication No. US20050130189A1
; GENERAL INFORMATION:
; APPLICANT: Paszicha, Pankaj
; APPLICANT: Shenoy, Mohan
; APPLICANT: Winston, John
; TITLE OF INVENTION: Compositions and Methods for Treating and Diagnosing
; FILE OF INVENTION: Irritable Bowel Syndrome
; FILE REFERENCE: 9511-136-27
; CURRENT APPLICATION NUMBER: US/10/923,035
; CURRENT FILING DATE: 2004-08-23
; PRIOR APPLICATION NUMBER: US 60/496,716
; PRIOR FILING DATE: 2003-08-21
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 967
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-923-035-37

Query Match 98.8%; Score 5224; DB 18; Length 967;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 956; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 1 MORAVPEGFGRKRLGSDMGNAERAPGSRSGFVPPTLLLLAAALLAVSDALGRPSEDEEL 60
Db 1 MORAVPEGFGRKRLGSDMGNAERAPGSRSGFVPPTLLLLAAALLAVSDALGRPSEDEEL 60

Qy 61 VPELERVPCHGTTTLRLHAFDQDLDPDSSFLAPGFTLQNVGRKSGSDTLPETDL 120
Db 61 VPELERVPCHGTTTLRLHAFDQDLDPDSSFLAPGFTLQNVGRKSGSETLPETDL 120

Qy 121 AHCFYSGTVNGDPSSAAALSCEGVGAFYLLGGEAYFIQPLPAASERLATAAPEKPPAP 180
Db 121 AHCFYSGTVNGDPSSAAALSCEGVGAFYLLGGEAYFIQPLPAASERLATAAPEKPPAP 180

Qy 181 LQFHLLRRNQDVGTCGVVDDEPRPTGKAETDEDEGEGEPQWSPDPAALQGVG 240
Db 181 LQFHLLRRNQDVGTCGVVDDEPRPTGKAETDEDEGEGEPQWSPDPAALQGVG 240

Qy 241 QPTGTGSIKRRKRVSSHRYVETMLVADQSMAPFHGSLKHVLLTLFVVAARLYKHPISRN 300
Db 241 QPTGTGSIKRRKRVSSHRYVETMLVADQSMAPFHGSLKHVLLTLFVVAARLYKHPISRN 300

Qy 301 SVSLVVVKILVIHDEQKGPVTSNAALTFLRNFQWQKHNPPSDRDAEHYDTAILFTRQD 360
```

Db 301 SVSLVVVKILVIHDEQKGPEVTSNAALTLRNFQWQKQHNPPSDRDAEHYDTAILFTROD 360
Qy 361 LCGSOTCDTLGWADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKOCASL 420
Db 361 LCGSOTCDTLGWADVGTCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKOCASL 420
Qy 421 NGVNQDSHMMASMLNLDHSQWSPSCGYMITSFLDNGHGECLMDKQNPQPIQLPGDLPGT 480
Db 421 NGVNQDSHMMASMLNLDHSQWSPSCGYMITSFLDNGHGECLMDKQNPQPIQLPGDLPGT 480
Qy 481 SYDANRQCQFTFGEDSKHCPCDAASTCTLWCTGTSGGVLCVOTKHPFWADGTSCEGKWC 540
Db 481 SYDANRQCQFTFGEDSKHCPCDAASTCTLWCTGTSGGVLCVOTKHPFWADGTSCEGKWC 540
Qy 541 INKGCVNKHRRKHFDTPFHSGWGMWGPWGDCSRTCGGGVQYTMRECDNPVPKNGGKYCEG 600
Db 541 INKGCVNKTRKHFDTPFHSGWGMWGPWGDCSRTCGGGVQYTMRECDNPVPKNGGKYCEG 600
Qy 601 KRVYRSCNLEDCPDNNGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKL 660
Db 601 KRVYRSCNLEDCPDNNGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKL 660
Qy 661 ICQAKGIGYFVLQPKVVDGTPCSPDSTSVCVQGCVKAGCDRIIDSKKKPKCGVCGGN 720
Db 661 ICQAKGIGYFVLQPKVVDGTPCSPDSTSVCVQGCVKAGCDRIIDSKKKPKCGVCGGN 720
Qy 721 GSTCKKISGVTSAKPGYHDIITPTGATNIEVKQNRQSRNNGSFLAIKAADGTIYL 780
Db 721 GSTCKKISGVTSAKPGYHDIITPTGATNIEVKQNRQSRNNGSFLAIKAADGTIYL 780
Qy 781 GDYTLSTLEQDIMYKGVWLYRSGSSAALERSFSPLKEPLTIQVLTGVNLRPKIKYTY 840
Db 781 GDYTLSTLEQDIMYKGVWLYRSGSSAALERSFSPLKEPLTIQVLTGVNLRPKIKYTY 840
Qy 841 FVKKKESFNALPTFSAWVIEBWGECSKSCELGWQRRLVECRDINGQPASECAKEVKPAS 900
Db 841 FVKKKESFNALPTFSAWVIEBWGECSKSCELGWQRRLVECRDINGQPASECAKEVKPAS 900
Qy 901 TRPCADHPCQWQWLGWSSCSKTCGKYKTSKLSHDDGVLSHSDCDPLKPKPHFIDF 960
Db 901 TRPCADHPCQWQWLGWSSCSKTCGKYKTSKLSHDDGVLSHSDCDPLKPKPHFIDF 960

RESULT 9

US-09-373-658-125
; Sequence 125, Application US/09373658
; Publication No. US20030092900A1
; GENERAL INFORMATION:
; APPLICANT: Iruela-Arispe, Luisa
; APPLICANT: Hastings, Gregg A.
; APPLICANT: Ruben, Steven M.
; APPLICANT: Jonak, Zdenka L.
; APPLICANT: Trulli, Stephen H.
; APPLICANT: Fronwald, James A.
; APPLICANT: Terrett, Jonathan A.
; TITLE OF INVENTION: Meth1 and Meth2 Polynucleotides and Polypeptides
; FILE REFERENCE: 1488.1070006
; CURRENT APPLICATION NUMBER: US/09/373,658
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 125
; LENGTH: 968
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-373-658-125

Query Match

98.8%; Score 5221.5; DB 10; Length 968;

Best Local Similarity 98.9%; Pred. No. 0;
Matches 957; Conservative 4; Mismatches 6; Indels 1; Gaps 1;
Qy 1 MORAVPEGFGRRKLGSD-MGNAERAPGSRSPGVPPTLLLLAAALAVSDALGRPSEDEE 59
Db 1 MORAVPEGFGRRKLGSDMMGNAERAPGSRSPGVPPTLLLLAAALAVSDALGRPSEDEE 60
Qy 60 LVVPELEVPVPGHTTTLRLHAFDQDLDLDPVPSDFSFLAPGFTLQNVGRKSGSDTLPETD 119
Db 61 LVVPELERAPGHTTTLRLHAFDQDLDLDPVPSDFSFLAPGFTLQNVGRKSGSETLPETD 120
Qy 120 LAHCFYSYTVNGDPSAAALSILCEGVRCGAFYLLGEAYFIQPLPAASERLATAAPGEKPPA 179
Db 121 LAHCFYSYTVNGDPSAAALSILCEGVRCGAFYLLGEAYFIQPLPAASERLATAAPGEKPPA 180
Qy 180 PLOFHLLLRNRQDVGVTGCVVVDDEPRPTGKAETDEDETEGEDEGEQWSPQDPALQGV 239
Db 181 PLOFHLLLRNRQDVGVTGCVVVDDEPRPTGKAETDEDETEGEDEGEQWSPQDPALQGV 240
Qy 240 GQPTGTGSIIRKRVFVSSHRYVETMLVADQSMAPFHGSLGHYLLTLFSAARLYKHPISIR 299
Db 241 GQPTGTGSIIRKRVFVSSHRYVETMLVADQSMAPFHGSLGHYLLTLFSAARLYKHPISIR 300
Qy 300 NSVSLVVVKILVIHDEQKGPEVTSNAALTLRNFQWQKQHNPPSDRDAEHYDTAILFTRO 359
Db 301 NSVSLVVVKILVIHDEQKGPEVTSNAALTLRNFQWQKQHNPPSDRDAEHYDTAILFTRO 360
Qy 360 DLCSGOTCDTLGMADVGTVPDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKOCAS 419
Db 361 DLCSGOTCDTLGMADVGTVPDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKOCAS 420
Qy 420 LNGVNQDSHMMASMLNLDHSQWSPSCGYMITSFLDNGHGECLMDKQNPQPIQLPGDLPG 479
Db 421 LNGVNQDSHMMASMLNLDHSQWSPSCGYMITSFLDNGHGECLMDKQNPQPIQLPGDLPG 480
Qy 480 TSYDANRQCQFTFGEDSKHCPCDAASTCTLWCTGTSGGVLCVOTKHPFWADGTSCEGKW 539
Db 481 TSYDANRQCQFTFGEDSKHCPCDAASTCTLWCTGTSGGVLCVOTKHPFWADGTSCEGKW 540
Qy 540 CINGKCVNKHRRKHFDTPFHSGWGMWGPWGDCSRTCGGGVQYTMRECDNPVPKNGGKYCE 599
Db 541 CINGKCVNKHRRKHFDTPFHSGWGMWGPWGDCSRTCGGGVQYTMRECDNPVPKNGGKYCE 600
Qy 600 KRVYRSCNLEDCPDNNGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCK 659
Db 601 KRVYRSCNLEDCPDNNGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCK 660
Qy 660 LIQAKGIGYFVLQPKVVDGTPCSPDSTSVCVQGCVKAGCDRIIDSKKKPKCGVCGG 719
Db 661 LIQAKGIGYFVLQPKVVDGTPCSPDSTSVCVQGCVKAGCDRIIDSKKKPKCGVCGG 720
Qy 720 NGSTCKKISGVTSAKPGYHDIITPTGATNIEVKQNRQSRNNGSFLAIKAADGTIYL 779
Db 721 NGSTCKKISGVTSAKPGYHDIITPTGATNIEVKQNRQSRNNGSFLAIKAADGTIYL 780
Qy 780 NGDYTLSTLEQDIMYKGVWLYRSGSSAALERSFSPLKEPLTIQVLTGVNLRPKIKYTY 839
Db 781 NGDYTLSTLEQDIMYKGVWLYRSGSSAALERSFSPLKEPLTIQVLTGVNLRPKIKYTY 840
Qy 840 YFVKKKESFNALPTFSAWVIEBWGECSKSCELGWQRRLVECRDINGQPASECAKEVKPA 899
Db 841 YFVKKKESFNALPTFSAWVIEBWGECSKSCELGWQRRLVECRDINGQPASECAKEVKPA 900
Qy 900 STRPCADHPCQWQWLGWSSCSKTCGKYKTSKLSHDDGVLSHSDCDPLKPKPHFID 959
Db 901 STRPCADHPCQWQWLGWSSCSKTCGKYKTSKLSHDDGVLSHSDCDPLKPKPHFID 960
Qy 960 FCTMAECS 967
Db 961 FCTMAECS 968

RESULT 10

```

US-09-373-658-2
; Sequence 2, Application US/09373658
; Publication No. US20030092900A1
; GENERAL INFORMATION:
; APPLICANT: Iruela-Arispe, Luisa
; APPLICANT: Hastings, Gregg A.
; APPLICANT: Ruben, Steven M.
; APPLICANT: Jonak, Zdenka L.
; APPLICANT: Trulli, Stephen H.
; APPLICANT: Fronwald, James A.
; APPLICANT: Terrett, Jonathan A.
; TITLE OF INVENTION: Meth1 and Meth2 Polynucleotides and Polypeptides
; FILE REFERENCE: 1488.1070006
; CURRENT APPLICATION NUMBER: US/09/373,658
; CURRENT FILING DATE: 1999-08-13
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 950
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-373-658-2

Query Match      97.3%; Score 5144; DB 10; Length 950;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 940; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 18 MGAERAPGSRSGFVPVPTLLLLAAALLAVSDALGRPSEDEELVPELVRPGHGTTLRL 77
DQ 1 MGAERAPGSRSGFVPVPTLLLLAAALLAVSDALGRPSEDEELVPELVRPGHGTTLRL 60
QY 78 LHAFDQQLDLVPPDSSFLAPGFTLQNVGRKSGSDTLPETDLAHCFYSGTVNGDPSSAA 137
DB 61 LHAFDQQLDLVPPDSSFLAPGFTLQNVGRKSGSETPLPETDLAHCFYSGTVNGDPSSAA 120
QY 138 ALSCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGKPPAPLQFHLLRRNRQGDVGGT 197
DB 121 ALSCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGKPPAPLQFHLLRRNRQGDVGGT 180
QY 198 CGVVDDDEPRPTGKAETDEDEGTEGEGPQSPDQALQGVQPTGTGSIKKRFVSSH 257
DB 181 CGVVDDDEPRPTGKAETDEDEGTEGEGPQSPDQALQGVQPTGTGSIKKRFVSSH 240
QY 258 RYVETMLVADQSMAEFHSGSLKHVLLTFLSVAARLYKHPSIRNSVSLVVKILVIHDEQ 317
DB 241 RYVETMLVADQSMAEFHSGSLKHVLLTFLSVAARLYKHPSIRNSVSLVVKILVIHDEQ 300
QY 318 GPEVTSNAALTNRNFCNWQKHNPPSDRDAEHYDTAILFTRQDLCSGOTCDTLGMADVGT 437
DB 361 VCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKQOCASLNGVNDSSHMASLSNL 420
QY 438 DHSQPWSPCSGYMITSFLDNGHGECLMDKPNQIQLPGDLPGTSYDANRQCQFTFGEDSK 497
DB 421 DHSQPWSPCSGYMITSFLDNGHGECLMDKPNQIQLPGDLPGTSYDANRQCQFTFGEDSK 480
QY 498 HCPDAAASTCSTLWCTGTSGGVLVCQTKHFPWADGTSCEGKWCINGKCNQKRNKRGHFDTP 557
DB 481 HCPDAAASTCSTLWCTGTSGGVLVCQTKHFPWADGTSCEGKWCINGKCNVKTDRKHFDTP 540
QY 558 FHSGWGMWGPWGDCSRTCCGGVQVYTMRECDNPVKNKGKCEGKRVYRSCNLEDCPDNN 617
DB 541 FHSGWGMWGPWGDCSRTCCGGVQVYTMRECDNPVKNKGKCEGKRVYRSCNLEDCPDNN 600
QY 618 KGTFREEQCEAHNBFASKAFSGGPAVEWIPKYAGVSPKDRCKLIQCAKGIGYFFVLQPKV 677
DB 601 KGTFREEQCEAHNBFASKAFSGGPAVEWIPKYAGVSPKDRCKLIQCAKGIGYFFVLQPKV 660
QY 678 VDGTPCSPDSTSVCCQGVKAGCDRIIDSKKKFKDGCVCGGNGSTCKKISGVS TSAKPG 737

```

```

DB 661 VDGTPCSPDSTSVCCQGVKAGCDRIIDSKKKFKDGCVCGGNGSTCKKISGVS TSAKPG 720
QY 738 YHDIITPTGATNLEVKORNGSRNNGSFIAIKAADGTYILNGDYTLSTLEQDLMYKGV 797
DB 721 YHDIITPTGATNLEVKORNGSRNNGSFIAIKAADGTYILNGDYTLSTLEQDLMYKGV 780
QY 798 VLRYSGSSAALERIRSFSPLEKPLTIQVLTGNALRPKIKTYTFVKKKESFNAIPTESA 857
DB 781 VLRYSGSSAALERIRSFSPLEKPLTIQVLTGNALRPKIKTYTFVKKKESFNAIPTESA 840
QY 858 WVIWEGECSKSCBLGWQRRLVECRDINGQPASCAKEVKPASTRCPADHPCPQWLGEW 917
DB 841 WVIWEGECSKSCBLGWQRRLVECRDINGQPASCAKEVKPASTRCPADHPCPQWLGEW 900
QY 918 SSCSKTCGKYKKTSLAKCLSHDGGVLSHSDCDPLKKPKHFDICTMAECS 967
DB 901 SSCSKTCGKYKKTSLAKCLSHDGGVLSHSDCDPLKKPKHFDICTMAECS 950

RESULT 11
US-09-989-687-2
; Sequence 2, Application US/09989687
; Publication No. US2004000249A1
; GENERAL INFORMATION:
; APPLICANT: Hastings, Gregg A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Meth1 and Meth2 Polynucleotides and Polypeptides
; FILE REFERENCE: 1488.107000D
; CURRENT APPLICATION NUMBER: US/09/989,687
; CURRENT FILING DATE: 2001-11-21
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 950
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-989-687-2

Query Match      97.3%; Score 5144; DB 11; Length 950;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 940; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 18 MGAERAPGSRSGFVPVPTLLLLAAALLAVSDALGRPSEDEELVPELVRPGHGTTLRL 77
DB 1 MGAERAPGSRSGFVPVPTLLLLAAALLAVSDALGRPSEDEELVPELVRPGHGTTLRL 60
QY 78 LHAFDQQLDLVPPDSSFLAPGFTLQNVGRKSGSDTLPETDLAHCFYSGTVNGDPSSAA 137
DB 61 LHAFDQQLDLVPPDSSFLAPGFTLQNVGRKSGSETPLPETDLAHCFYSGTVNGDPSSAA 120
QY 138 ALSCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGKPPAPLQFHLLRRNRQGDVGGT 197
DB 121 ALSCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGKPPAPLQFHLLRRNRQGDVGGT 180
QY 198 CGVVDDDEPRPTGKAETDEDEGTEGEGPQSPDQALQGVQPTGTGSIKKRFVSSH 257
DB 181 CGVVDDDEPRPTGKAETDEDEGTEGEGPQSPDQALQGVQPTGTGSIKKRFVSSH 240
QY 258 RYVETMLVADQSMAEFHSGSLKHVLLTFLSVAARLYKHPSIRNSVSLVVKILVIHDEQ 317
DB 241 RYVETMLVADQSMAEFHSGSLKHVLLTFLSVAARLYKHPSIRNSVSLVVKILVIHDEQ 300
QY 318 GPEVTSNAALTNRNFCNWQKHNPPSDRDAEHYDTAILFTRQDLCSGOTCDTLGMADVGT 377
DB 301 GPEVTSNAALTNRNFCNWQKHNPPSDRDAEHYDTAILFTRQDLCSGOTCDTLGMADVGT 360
QY 378 VCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKQOCASLNGVNDSSHMASLSNL 437
DB 361 VCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKQOCASLNGVNDSSHMASLSNL 420
QY 438 DHSQPWSPCSGYMITSFLDNGHGECLMDKPNQIQLPGDLPGTSYDANRQCQFTFGEDSK 497
DB 421 DHSQPWSPCSGYMITSFLDNGHGECLMDKPNQIQLPGDLPGTSYDANRQCQFTFGEDSK 480

```

QY 498 HCPDAASTCSTLWCTGTSGVLCVQTKHFPWADGTSCEGKWCINGKCNKHKHFDTP 557
DB 481 HCPDAASTCSTLWCTGTSGVLCVQTKHFPWADGTSCEGKWCINGKCNKHKHFDTP 540
QY 558 FHGSWMGMPWDCSRTCGGGVQYTMRECDNVPKNGGKYCEGKVRVRSCHLEDCPDNN 617
DB 541 FHGSWMGMPWDCSRTCGGGVQYTMRECDNVPKNGGKYCEGKVRVRSCHLEDCPDNN 600
QY 618 GXTFREEOCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLI COAKGIGYFFVLQPKV 677
DB 601 GXTFREEOCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLI COAKGIGYFFVLQPKV 660
QY 678 VDGTPCSPDSTSVQVQGCVKAGCDRIIDSKKFKDKGCVGNGSTCKKISGSVTSAPKG 737
DB 661 VDGTPCSPDSTSVQVQGCVKAGCDRIIDSKKFKDKGCVGNGSTCKKISGSVTSAPKG 720
QY 738 YHDIITPTGATNIEVKQNRQSRNNGSFLAIKAADGTYYILNGDYTLSTLEQDIMYKGV 797
DB 721 YHDIITPTGATNIEVKQNRQSRNNGSFLAIKAADGTYYILNGDYTLSTLEQDIMYKGV 780
QY 798 VLRYSGSSAALERIRSFSPKLEPLTIQVLTUGNALRPKIKYTYFVKKKKESFNAIPTFSA 857
DB 781 VLRYSGSSAALERIRSFSPKLEPLTIQVLTUGNALRPKIKYTYFVKKKKESFNAIPTFSA 840
QY 858 WVEEWGECSSCELGWQRRLECDINGQPASECAKEVKPASTRPCADHPCPQWLGEW 917
DB 841 WVEEWGECSSCELGWQRRLECDINGQPASECAKEVKPASTRPCADHPCPQWLGEW 900
QY 918 SSCSKTCGKYKTSLSKCLSHDGGVLSHSDCPLKKPKHFIDFCTMAECS 967
DB 901 SSCSKTCGKYKTSLSKCLSHDGGVLSHSDCPLKKPKHFIDFCTMAECS 950

RESULT 12

US-10-667-281-2
; Sequence 281, Application US/10667281
; Publication No. US20050100916A1
; GENERAL INFORMATION:
; APPLICANT: Klonowski, Paul
; APPLICANT: Allard, John
; APPLICANT: Heller, Renu
; APPLICANT: Van Wart, Harold
; TITLE OF INVENTION: Human Aggreganase and Nucleic Acid
; TITLE OF INVENTION: Compositions Encoding the Same
; FILE REFERENCE: ROCH-002
; CURRENT APPLICATION NUMBER: US/10/667,281
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: US/09/568,559
; PRIOR FILING DATE: 2000-05-09
; PRIOR APPLICATION NUMBER: 60/133,343
; PRIOR FILING DATE: 1999-05-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 949
; TYPE: PRT
; ORGANISM: human
US-10-667-281-2

Query Match 97.2%; Score 5140; DB 17; Length 949;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 939; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 19 GNAERAPGSRSGPVTLLILAAALLAVSDALGRSEDEELVPELERVPGHGTURL 78
DB 1 GNAERAPGSRSGPVTLLILAAALLAVSDALGRSEDEELVPELERVPGHGTURL 60
QY 79 HAFDQOLDLVDPPDSFLAPGFTLQWGRKSGSDTPLPETDLAHCFYSGTVNGDPSAAA 138
DB 61 HAFDQOLDLVDPPDSFLAPGFTLQWGRKSGSDTPLPETDLAHCFYSGTVNGDPSAAA 120
QY 139 LSLCEGVRGAFYLLGEAYFIQPLPAASERLATAAPEKPPAPLQPHLLRRNRQGDVGTC 198

DB 121 LSLCEGVRGAFYLLGEAYFIQPLPAASERLATAAPEKPPAPLQPHLLRRNRQGDVGTC 180
QY 199 GVDDDEPRPTGKAETEDDEGTEGEGPQWSPDPAQGVGQPTGTGSIKKRKRFSVSHR 258
DB 181 GVDDDEPRPTGKAETEDDEGTEGEGPQWSPDPAQGVGQPTGTGSIKKRKRFSVSHR 240
QY 259 YVETMLVADQMAEFHSGSLKHYLLTLFSVAARLYKHPISIRNSVSLVVKILVHDEQK 318
DB 241 YVETMLVADQMAEFHSGSLKHYLLTLFSVAARLYKHPISIRNSVSLVVKILVHDEQK 300
QY 319 PEVTSNAALTIRNFCNMQHNPSPDRDAEHYDTAILFTRQDLCSQTCDTLGMADVGV 378
DB 301 PEVTSNAALTIRNFCNMQHNPSPDRDAEHYDTAILFTRQDLCSQTCDTLGMADVGV 360
QY 379 CDPSSCSVIEDDGLQAAFTTAHELGHVFNPNPHDDAKOCASLNGVQDSHMMASLSNLD 438
DB 361 CDPSSCSVIEDDGLQAAFTTAHELGHVFNPNPHDDAKOCASLNGVQDSHMMASLSNLD 420
QY 439 HSQWSPCSGYMITSFLDNHGECLMDKPNPIQLPGDLPGTSYDANRQCOFTFGEDESKH 498
DB 421 HSQWSPCSGYMITSFLDNHGECLMDKPNPIQLPGDLPGTSYDANRQCOFTFGEDESKH 480
QY 499 CPDAASTCSTLWCTGTSGVLCVQTKHFPWADGTSCEGKWCINGKCNKHKHFDTP 558
DB 481 CPDAASTCSTLWCTGTSGVLCVQTKHFPWADGTSCEGKWCINGKCNKHKHFDTP 540
QY 559 HGSWMGMPWDCSRTCGGGVQYTMRECDNVPKNGGKYCEGKVRVRSCHLEDCPDNN 618
DB 541 HGSWMGMPWDCSRTCGGGVQYTMRECDNVPKNGGKYCEGKVRVRSCHLEDCPDNN 600
QY 619 KTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLI COAKGIGYFFVLQPKV 678
DB 601 KTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLI COAKGIGYFFVLQPKV 660
QY 679 DGTFCSPDSTSVQVQGCVKAGCDRIIDSKKFKDKGCVGNGSTCKKISGSVTSAPKG 738
DB 661 DGTFCSPDSTSVQVQGCVKAGCDRIIDSKKFKDKGCVGNGSTCKKISGSVTSAPKG 720
QY 739 HDIITPTGATNIEVKQNRQSRNNGSFLAIKAADGTYYILNGDYTLSTLEQDIMYKGV 798
DB 721 HDIITPTGATNIEVKQNRQSRNNGSFLAIKAADGTYYILNGDYTLSTLEQDIMYKGV 780
QY 799 LRYSGSSAALERIRSFSPKLEPLTIQVLTUGNALRPKIKYTYFVKKKKESFNAIPTFSA 858
DB 781 LRYSGSSAALERIRSFSPKLEPLTIQVLTUGNALRPKIKYTYFVKKKKESFNAIPTFSA 840
QY 859 VIEEWGECSSCELGWQRRLECDINGQPASECAKEVKPASTRPCADHPCPQWLGEW 918
DB 841 VIEEWGECSSCELGWQRRLECDINGQPASECAKEVKPASTRPCADHPCPQWLGEW 900
QY 919 SSCSKTCGKYKTSLSKCLSHDGGVLSHSDCPLKKPKHFIDFCTMAECS 967
DB 901 SSCSKTCGKYKTSLSKCLSHDGGVLSHSDCPLKKPKHFIDFCTMAECS 949

RESULT 13

US-10-741-600-1605
; Sequence 1605, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1605
; LENGTH: 950
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-741-600-1605

```

Query Match      97.1%; Score 5136; DB 17; Length 950;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 939; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 18 MGNAERAPGSRSGFPGVPTLLALLAAALAVSDALGRPSEDEELVPELVRVPHGHTTLRL 77
DB 1 MGNAERAPGSRSGFPGVPTLLALLAAALAVSDALGRPSEDEELVPELVRVPHGHTTLRL 60

QY 78 LHAFDQQLDLVPPDSSFLAPGFTLQNVGRKSGSDTLPETDLAHCFFYSGTVNGDPSSAA 137
DB 61 LHAFDQQLDLVPPDSSFLAPGFTLQNVGRKSGSDTLPETDLAHCFFYSGTVNGDPSSAA 120

QY 138 ALSLCEGVGAFYLLGEAYFIQPLPAASERLATAAPGKPPAPLOPHLLRRNRQDVGVT 197
DB 121 ALSLCEGVGAFYLLGEAYFIQPLPAASERLATAAPGKPPAPLOPHLLRRNRQDVGVT 180

QY 198 CGVVDDEPRPTGKAETDEDEGTEGEDEGQWSPDPALQGVGQPTGTGSIKRRKRVSSH 257
DB 181 CGVVDDEPRPTGKAETDEDEGTEGEDEGQWSPDPALQGVGQPTGTGSIKRRKRVSSH 240

QY 258 RYVETMLVADQSMABFHSGGLKHYLLTLFVSVAARLYKHPSIRNSVSLVVKILVIHDEQK 317
DB 241 RYVETMLVADQSMABFHSGGLKHYLLTLFVSVAARLYKHPSIRNSVSLVVKILVIHDEQK 300

QY 318 GPEVTSNAALTURNCNWKQHNPPSDRDAEHYDTAILFTRODLCSGQTCOTLGMADVGT 377
DB 301 GPEVTSNAALTURNCNWKQHNPPSDRDAEHYDTAILFTRODLCSGQTCOTLGMADVGT 360

QY 378 VCDPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACSLNGVNDQSHMMASMLSNL 437
DB 361 VCDPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACSLNGVNDQSHMMASMLSNL 420

QY 438 DHSQWSPSCGYWITSTFLDNGHGECLMDKQNPQIQLPGDLPGTSYDANRQCOFTFGEDSK 497
DB 421 DHSQWSPSCGYWITSTFLDNGHGECLMDKQNPQIQLPGDLPGTSYDANRQCOFTFGEDSK 480

QY 498 HCPDAASTCTLWCTGTSGGVLVCOTKHFPWADGTSCEGKWCINGKCNVNRKHGHTDP 557
DB 481 HCPDAASTCTLWCTGTSGGVLVCOTKHFPWADGTSCEGKWCINGKCNVNRKHGHTDP 540

QY 558 FHGSGWGMGPWGDSCSRTCGGVQYTMRECDNPVPKNGGKCEGKRVYRSCNLEDCPDNN 617
DB 541 FHGSGWGMGPWGDSCSRTCGGVQYTMRECDNPVPKNGGKCEGKRVYRSCNLEDCPDNN 600

QY 618 GKTFRBEOCEAHNEFSAFSGGPAVEWIPKYAGVSPKDRCKLICQAKGIGYFFVLQPKV 677
DB 601 GKTFRBEOCEAHNEFSAFSGGPAVEWIPKYAGVSPKDRCKLICQAKGIGYFFVLQPKV 660

QY 678 VDGTPCSPDSTSVCOGQCVKAGCDRIIDSKKKFKDCKGCGNGSTCKKISGSVTSAPKG 737
DB 661 VDGTPCSPDSTSVCOGQCVKAGCDRIIDSKKKFKDCKGCGNGSTCKKISGSVTSAPKG 720

QY 738 YHDITITPTGATNIEVKQNRGSRNNGSFLAIIKAADGTIILNGDYTLSTLEQIMYKGV 797
DB 721 YHDITITPTGATNIEVKQNRGSRNNGSFLAIIKAADGTIILNGDYTLSTLEQIMYKGV 780

QY 798 VLRYSGSSAALERIRSPFLKEPLTIQVLTVGNALRPKIKYTVFKKKKESFNAIPTPSA 857
DB 781 VLRYSGSSAALERIRSPFLKEPLTIQVLTVGNALRPKIKYTVFKKKKESFNAIPTPSA 840

QY 858 WVIEBWGCSKSELGWORRLVECRDINGQPASECAKEVKPASTRCPADHPCPQWOLGEW 917
DB 841 WVIEBWGCSKSELGWORRLVECRDINGQPASECAKEVKPASTRCPADHPCPQWOLGEW 900

QY 918 SSCSKTCGKYKTSKLSCLSHDGGVLSHSDCDPLKKPKHFIDFCTMAECS 967
DB 901 SSCSKTCGKYKTSKLSCLSHDGGVLSHSDCDPLKKPKHFIDFCTMAECS 950

```

RESULT 14

US-10-973-858-14

```

; Sequence 14, Application US/10973858
; Publication No. US20050176030A1
; GENERAL INFORMATION:
; APPLICANT: Gan, Li
; APPLICANT: Gonzalez-Zulueta, Mirella
; APPLICANT: Ye, Shiming
; APPLICANT: Urfer, Roman
; APPLICANT: Nickolich, Karoly
; TITLE OF INVENTION: Regulated Nucleic Acids in Pathogenesis of Alzheimer's Disease
; FILE REFERENCE: AGYT-047
; CURRENT APPLICATION NUMBER: US/10/973,858
; CURRENT FILING DATE: 2004-10-25
; PRIOR APPLICATION NUMBER: 60/515,562
; PRIOR FILING DATE: 2003-10-28
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 14
; LENGTH: 950
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-973-858-14

```

```

Query Match      97.1%; Score 5136; DB 18; Length 950;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 939; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

```

```

QY 18 MGNAERAPGSRSGFPGVPTLLALLAAALAVSDALGRPSEDEELVPELVRVPHGHTTLRL 77
DB 1 MGNAERAPGSRSGFPGVPTLLALLAAALAVSDALGRPSEDEELVPELVRVPHGHTTLRL 60

QY 78 LHAFDQQLDLVPPDSSFLAPGFTLQNVGRKSGSDTLPETDLAHCFFYSGTVNGDPSSAA 137
DB 61 LHAFDQQLDLVPPDSSFLAPGFTLQNVGRKSGSDTLPETDLAHCFFYSGTVNGDPSSAA 120

QY 138 ALSLCEGVGAFYLLGEAYFIQPLPAASERLATAAPGKPPAPLOPHLLRRNRQDVGVT 197
DB 121 ALSLCEGVGAFYLLGEAYFIQPLPAASERLATAAPGKPPAPLOPHLLRRNRQDVGVT 180

QY 198 CGVVDDEPRPTGKAETDEDEGTEGEDEGQWSPDPALQGVGQPTGTGSIKRRKRVSSH 257
DB 181 CGVVDDEPRPTGKAETDEDEGTEGEDEGQWSPDPALQGVGQPTGTGSIKRRKRVSSH 240

QY 258 RYVETMLVADQSMABFHSGGLKHYLLTLFVSVAARLYKHPSIRNSVSLVVKILVIHDEQK 317
DB 241 RYVETMLVADQSMABFHSGGLKHYLLTLFVSVAARLYKHPSIRNSVSLVVKILVIHDEQK 300

QY 318 GPEVTSNAALTURNCNWKQHNPPSDRDAEHYDTAILFTRODLCSGQTCOTLGMADVGT 377
DB 301 GPEVTSNAALTURNCNWKQHNPPSDRDAEHYDTAILFTRODLCSGQTCOTLGMADVGT 360

QY 378 VCDPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACSLNGVNDQSHMMASMLSNL 437
DB 361 VCDPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHDDAKQACSLNGVNDQSHMMASMLSNL 420

QY 438 DHSQWSPSCGYWITSTFLDNGHGECLMDKQNPQIQLPGDLPGTSYDANRQCOFTFGEDSK 497
DB 421 DHSQWSPSCGYWITSTFLDNGHGECLMDKQNPQIQLPGDLPGTSYDANRQCOFTFGEDSK 480

QY 498 HCPDAASTCTLWCTGTSGGVLVCOTKHFPWADGTSCEGKWCINGKCNVNRKHGHTDP 557
DB 481 HCPDAASTCTLWCTGTSGGVLVCOTKHFPWADGTSCEGKWCINGKCNVNRKHGHTDP 540

QY 558 FHGSGWGMGPWGDSCSRTCGGVQYTMRECDNPVPKNGGKCEGKRVYRSCNLEDCPDNN 617
DB 541 FHGSGWGMGPWGDSCSRTCGGVQYTMRECDNPVPKNGGKCEGKRVYRSCNLEDCPDNN 600

QY 618 GKTFRBEOCEAHNEFSAFSGGPAVEWIPKYAGVSPKDRCKLICQAKGIGYFFVLQPKV 677
DB 601 GKTFRBEOCEAHNEFSAFSGGPAVEWIPKYAGVSPKDRCKLICQAKGIGYFFVLQPKV 660

QY 678 VDGTPCSPDSTSVCOGQCVKAGCDRIIDSKKKFKDCKGCGNGSTCKKISGSVTSAPKG 737
DB 661 VDGTPCSPDSTSVCOGQCVKAGCDRIIDSKKKFKDCKGCGNGSTCKKISGSVTSAPKG 720

```



```
QY 738 YHDIITPTGATNIEVKQNRGRNNGSFLAIKAADGTYILNGDYTLSTLEQDIMYKGV 797
Db 721 YHDIITPTGATNIEVKQNRGRNNGSFLAIKAADGTYILNGDYTLSTLEQDIMYKGV 780
QY 798 VLRYGSSAALIRIRSFSLKEPLTIQVLTGNALRPKIKYTFVKKKESFNAIPTFSA 857
Db 781 VLRYGSSAALIRIRSFSLKEPLTIQVLTGNALRPKIKYTFVKKKESFNAIPTFSA 840
QY 858 WYIEWGECSKCELGWQRRLVECRDINGQPASECAKEVKPASTRPCADHPCPQWOLGEW 917
Db 841 WYIEWGECSKCELGWQRRLVECRDINGQPASECAKEVKPASTRPCADHPCPQWOLGEW 900
QY 918 SSCSKTCGKGYYKTSILKCLSHDGGVLSHSDCLPKPKPHFIDFCTMAECS 967
Db 901 SSCSKTCGKGYYKTSILKCLSHDGGVLSHSDCLPKPKPHFIDFCTMAECS 950

RESULT 15
US-09-741-151-4
; Sequence 4, Application US/09741151
; Publication No. US20020086400A1
; GENERAL INFORMATION:
; APPLICANT: ZHU, Shaoping et al
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CL001005
; CURRENT APPLICATION NUMBER: US/09/741,151
; CURRENT FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 931
; TYPE: PRT
; ORGANISM: Human
US-09-741-151-4

Query Match 95.3%; Score 5036; DB 9; Length 931;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 920; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 36 LLLLLAALLAVSDALGRPSEDEELVPELERVPGHGTTRLRHLHAFDQQLDLDPDSSF 95
Db 1 LLLLLAALLAVSDALGRPSEDEELVPELERVPGHGTTRLRHLHAFDQQLDLDPDSSF 60
QY 96 LAPGFTLQNVGRKSGDTPLPETDLAHCFYGTNGDPSSAALSLCEGVRGAFYLLGEA 155
Db 61 LAPGFTLQNVGRKSGSETPLPETDLAHCFYGTNGDPSSAALSLCEGVRGAFYLLGEA 120
QY 156 YFIQPLPAASERLATAAPGKEPPAPLQFHLLRRNRQGVGGTCGVVDDPRPTGKAETED 215
Db 121 YFIQPLPAASERLATAAPGKEPPAPLQFHLLRRNRQGVGGTCGVVDDPRPTGKAETED 180
QY 216 EDEGTEGEDEGPQSPQDPAQGVGQPTGTGIRKRFVSSHRYVETMLVADQSMAPFHG 275
Db 181 EDEGTEGEDEGAQSPQDPAQGVGQPTGTGIRKRFVSSHRYVETMLVADQSMAPFHG 240
QY 276 SGLKHYYLLTFSVAARLYKHPISIRNSVLSLVVKILVHDEQGPVTSNAAITLRNFCNW 335
Db 241 SGLKHYYLLTFSVAARLYKHPISIRNSVLSLVVKILVHDEQGPVTSNAAITLRNFCNW 300
QY 336 QKQHNPPSDRDAEHYDTALTFRQDLCSGQTCDTLGMADVGTCDPSPRSCSVIEDDGLQA 395
Db 301 QKQHNPPSDRDAEHYDTALTFRQDLCSGQTCDTLGMADVGTCDPSPRSCSVIEDDGLQA 360
QY 396 APTTAHELGHVFNMPHDDAKQCASLNGVNDQSHMMASMLSNLDHQPWSPSCGYMITSF 455
Db 361 APTTAHELGHVFNMPHDDAKQCASLNGVNDQSHMMASMLSNLDHQPWSPSCGYMITSF 420
QY 456 DNGHGECLMDKQNPQLPGDLPGTISYDANROQCTTFGSDSKHCPDAASTCTLWCTGTS 515
Db 421 DNGHGECLMDKQNPQLPGDLPGTISYDANROQCTTFGSDSKHCPDAASTCTLWCTGTS 480
```

```
QY 516 GGVLVQCQTKHPFPWADGTSCEGKWCINGKCVNKNRKHFDTPFHGSMGMWGPWGDCSRTC 575
Db 481 GGVLVQCQTKHPFPWADGTSCEGKWCINGKCVNKNRKHFDTPFHGSMGMWGPWGDCSRTC 540
QY 576 GGGVQYTNRECDNPVPKNGGKYCEGKRVYRSCNLEDCPDNNGKTFREEOCEAHNEFSKA 635
Db 541 GGGVQYTNRECDNPVPKNGGKYCEGKRVYRSCNLEDCPDNNGKTFREEOCEAHNEFSKA 600
QY 636 SFGSGPAVEWIPKYAGVSPKDRCKLI COAKGIGYFFVLQPKVVDGTPCSPDSTSVCVQGG 695
Db 601 SFGSGPAVEWIPKYAGVSPKDRCKLI COAKGIGYFFVLQPKVVDGTPCSPDSTSVCVQGG 660
QY 696 CVKAGCDRIIDSXXXKFKCGVCGGNGSTCKKISGSVTSAPKPGYHDIITPTGATNIEVKQ 755
Db 661 CVKAGCDRIIDSXXXKFKCGVCGGNGSTCKKISGSVTSAPKPGYHDIITPTGATNIEVKQ 720
QY 756 RNQGRSRNNGSFLAIKAADGTYILNGDYTLSTLBODIMYKGVLRYSGSSAALERIRSPS 815
Db 721 RNQGRSRNNGSFLAIKAADGTYILNGDYTLSTLBODIMYKGVLRYSGSSAALERIRSPS 780
QY 816 PLKEPLTIQVLTGNALRPKIKYTFVKKKESFNAIPTFSAWIEEWGECSSKCELGWQ 875
Db 781 PLKEPLTIQVLTGNALRPKIKYTFVKKKESFNAIPTFSAWIEEWGECSSKCELGWQ 840
QY 876 RRLVECRDINGQPASECAKEVKPASTRPCADHPCPQWOLGEWSSCSKTCGKGYKTSKLC 935
Db 841 RRLVECRDINGQPASECAKEVKPASTRPCADHPCPQWOLGEWSSCSKTCGKGYKTSKLC 900
QY 936 LSHDGGVLSHSDCLPKPKPHFIDFCTMAEC 966
Db 901 LSHDGGVLSHSDCLPKPKPHFIDFCTMAEC 931
```

Search completed: August 22, 2005, 10:23:20
Job time : 171 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2005, 10:09:09 ; Search time 47 Seconds
(without alignments)
1979.609 Million cell updates/sec

Title: US-09-989-687-126

Perfect score: 5287
Sequence: 1 MRAVPEGFGRKRLGSDMGN.....CDPLKKPKHFIDFCTMAECS 967

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 9621673 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4264.5	80.7	951	2 T00017	Gene ADAMTS-1 prot
2	3056	57.8	550	2 T47158	hypothetical prote
3	2129	40.3	837	2 T00355	hypothetical prote
4	1598	30.2	2165	2 T21371	hypothetical prote
5	1096	20.7	1205	2 T18517	procollagen N-endo
6	645	12.2	1558	2 C89114	protein C37C3.6a [
7	645	12.2	2167	2 T34395	hypothetical prote
8	614.5	11.6	1444	2 T18856	angiogenesis inhib
9	466	8.8	860	2 T18892	hypothetical prote
10	418.5	7.9	1059	2 T22545	hypothetical prote
11	402	7.6	957	2 T15976	hypothetical prote
12	391	7.4	951	2 T00260	hypothetical prote
13	378.5	7.2	903	2 S60257	metrin alpha - mo
14	347	6.6	549	2 S48169	metalloproteinase
15	343	6.5	571	2 S24789	Jararhagin C.precu
16	343	6.5	617	2 S48160	metalloproteinase
17	340.5	6.4	610	2 JC8056	halysase - Glycidu
18	326.5	6.2	610	2 JC7530	vascular apoptosis
19	326.5	6.2	826	2 A60385	monocyte surface a
20	325.5	6.2	609	2 S55270	catrocollastatin p
21	321	6.1	655	2 JC7850	disintegrin and me
22	320.5	6.1	481	2 JC4342	fibrinolytic prote
23	308.5	5.8	481	2 S43125	trilmucin precursor
24	304.5	5.8	407	2 S63260	metalloproteinase
25	304	5.7	411	1 HV5NFA	fibrolase (EC 3.4.
26	302	5.7	616	2 A55796	ecarin precursor -
27	301	5.7	478	2 A43296	atrolysin E (EC 3.
28	299	5.7	478	2 JQ1301	hemorrhagic protei
29	296.5	5.6	478	2 JC4880	fibrinolytic metal

30	295.5	5.6	814	2	G02390	disintegrin-like m
31	287.5	5.4	480	1	A30065	trigramin precursor
32	277.5	5.2	789	2	S28259	androgen-regulated
33	276.5	5.2	484	2	JC8020	metalloproteinase-
34	276	5.2	1572	2	T00027	brain-specific ang
35	272	5.1	660	2	S71949	metalloproteinase
36	271	5.1	414	1	HYRSAC	atrolysin C (EC 3.
37	271	5.1	414	2	S41609	atrolysin C (EC 3.
38	268	5.1	952	2	T18900	disintegrin and me
39	267.5	5.1	1170	2	A40558	thrombospondin 1 p
40	265.5	5.0	508	2	T22836	hypothetical prote
41	264	5.0	1074	2	JC5928	semaphorin F presu
42	260	4.9	776	2	S28258	androgen-regulated
43	257.5	4.9	419	2	S41607	atrolysin A (EC 3.
44	256	4.8	1170	1	TSHUP1	thrombospondin 1 p
45	254.5	4.8	670	2	I65967	disintegrin-like m

ALIGNMENTS

RESULT 1

T00017
Gene ADAMTS-1 protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 15-Mar-2004
C:Accession: T00017
R:Kuno, K.; Lizasa, H.; Ohno, S.; Matsushima, K.
Genomics 46, 466-471, 1997
A>Title: The exon/intron organization and chromosomal mapping of the mouse ADAMTS-1 gene
A:Reference number: Z14055; MUID:98110583; PMID:9441751
A:Accession: T00017
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-951 <KUN>
A:Cross-references: EMBL:AB001735; NID:g2809056; PIDN:BAA24501.1; PID:g2809057
A:Experimental source: strain 129SVJ
C:Genetics:
A:Gene: ADAMTS-1
A:Introns: 228/1; 343/3; 388/1; 444/1; 539/3; 602/1; 660/3; 719/2
F:542-598/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match	80.7%;	Score 4264.5;	DB 2;	Length 951;
Best Local Similarity	81.6%;	Pred. No. 2.6e-278;		
Matches	782;	Conservative	56;	Mismatches 105; Indels 15; Gaps 5;
Qy	18	MGNARAPGSRFGVPVTLILLAAA---LLAVSDALGRPSEDEELVPELERYVGHG--	72	
Db	1	MGDVQRAARSRSGLSAHMLLLLLASITWLLCARGAHRPTDEDEELVLSLERAPGHDS	60	
Qy	73	TTRLRLHAFDOQLDLDPDSSFLAPGFTLQNVGRKSGSDTPL--PETDLAHCYSGTVN	130	
Db	61	TTRLRLDAFGQQLHLKLQDPSGFLAPGFTLTGVRSGPSEAQLDPTGDLAHCYSGTVN	120	
Qy	131	GDPSAAALSICEGVRGAFYLLGEAYFTQPLP-AASERLATAAPEKPPAPLQPHLLRRN	189	
Db	121	GDPSAAALSICEGVRGAFYLLQGEFFIQAPGAVATERLAPAPPEESSARPPQHILRRR	180	
Qy	190	RQGDVGTCGVVDDEPRPTGKAETDEDEGTEGEDEGPQSPDPALQGVQPTGTSIR	249	
Db	181	RRGGGAKCGVMDDETLLP-----SDSRPESQNTQNPVRDPTPDADKPGSGSIR	233	
Qy	250	KKRFVSSHRYVETMLVDQSMAEFGSLKHYLLTLFSVAARLYKHPISIRNSVLVVVKI	309	
Db	234	KKRFVSSPRYVETMLVDQSMADFGSLKHYLLTLFSVAARFYKHPISIRNSISLVVKI	293	
Qy	310	LVIHDEQKGPVTSNAALTTLRNFCNQKHNPSPDRDDEHYDTAILFTRQDLCSGQTCOT	369	
Db	294	LVIIIEQKGPVTSNAALTTLRNFCNQKHNPSPDRDDEHYDTAILFTRQDLCSGHTCDT	353	
Qy	370	LGMDVGVTCVDPSPSCSVIEDDGLQAATTAAHELGHVFNMPHDDAKKQCASLNGVNDSHM	429	
Db	354	LGMDVGVTCVDPSPSCSVIEDDGLQAATTAAHELGHVFNMPHDDAKKCHCASLNGVTGDSHL	413	

```
QY 430 MASMLSNLDHSQPSPCSGYMLTSLFDNGHGECIMDKPQNPIQLPGDLPGTSYDANROCO 489
Db 414 MASMLSSLDHSPSPCSAYMTSLFDNGHGECIMDKPQNPIQLPSDLPGLTYDANROCO 473
QY 490 FTFGSDSKHCPDAASTCTSLWCTGTSGGVLVCQTQKHPWADGTSCEGKWCINGKCVNKN 549
Db 474 FTFGESKHCPDAASTCTTLWCTGTSGGLLVCQTQKHPWADGTSCEGKWCVSGKCVNKT 533
QY 550 HKKHFDTPFHGSGMWGPGWDCSRTCCTGGGVQVYTMRECDNPVPKNGGKYCEGKRVYRSCN 609
Db 534 DMKHEPATPVHSGWPGWPGWDCSRTCCTGGGVQVYTMRECDNPVPKNGGKYCEGKRVYRSCN 593
QY 610 LEDCDNNNGKTFREOCEAHNEFSKASFGSPGPAVEMIPKYAGVSPKDRCKLICQAKGIGY 669
Db 594 IEDCDNNNGKTFREOCEAHNEFSKASFGNEFTVEWTPKYAGVSPKDRCKLTCEAKGIGY 653
QY 670 FVLQPKVVDGTPPCSPDSTSVCVQGCVKAGCDRIIDSKKFDPKCGVCGGNGSTCKKISG 729
Db 654 FVLQPKVVDGTPPCSPDSTSVCVQGCVKAGCDRIIDSKKFDPKCGVCGGNGSTCKKMSG 713
QY 730 SVTSAPGVDHDIITPTGATNIEVKQNRGRNNGSFLAIKAADGTYILNGDYTLSTLE 789
Db 714 IVTSTRPGVDHDIITIPAGATNIEVKHRNQRGRNNGSFLAIRAADGTYILNGNFTLSTLE 773
QY 790 QDIMYKGVVLRYSYGSSAALERIRSFPLKEPTIOVLTVGNALRPKIYTYFVKKKESF 849
Db 774 QDLTYKGTVLRYSYGSSAALERIRSFPLKEPTIOVLTVGNALRPKIYTYFVKKKTESF 833
QY 850 NAIPTFSAWVIEEWGECSSKBLGQRRRLVECRDINGOPASECAKEVPASTRPCADHP 909
Db 834 NAIPTFSEWVIEEWGECSTCGSGHQRRVVQCRDINGHPASECAKEVPASTRPCADLP 893
QY 910 PQWLQGEWSSCKTCGKGYKTSLSKLSHGGVLSHSDCDPLKPKPHIDFCTMAECS 967
Db 894 PHQVQGDWSPCKTCGKGYKRTKLCVSHDGGVLSNESCDPLKPKPHYDFCTLTQCS 951

RESULT 2
T47158
hypothetical protein DKFp762C110.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C:Accession: T47158
R:Blum, H.; Baurisachs, S.; Mewes, H.W.; Weil, B.; Wiemann, S.
submitted to the Protein Sequence Database, March 2000
A:Reference number: Z24379
A:Accession: T47158
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-550 <AAA>
A:Cross-references: UNIPROT:Q9UHI8; EMBL:AL162080
A:Experimental source: adult melanoma (Mewo cell line); clone DKFzp762C110
C:Genetics:
A:Note: DKFzp762C110.1

Query Match 57.8%; Score 3056; DB 2; Length 550;
Best Local Similarity 99.1%; Pred. No. 1.9e-197;
Matches 545; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 418 ASLNGVNDSHMWSMLNLDHSQPSPCSGYMTSFLDNGHGECIMDKPQNPIQLPGDL 477
Db 1 ASLNGVNDSHMWSMLNLDHSQPSPCSAYMTSFLDNGHGECIMDKPQNPIQLPGDL 60
QY 478 POTSVDANRQCOFTFGEDSKHCPDAASTCTSLWCTGTSGGVLVCQTKHPWADGTSCEG 537
Db 61 POTSVDANRQCOFTFGEDSKHCPDAASTCTSLWCTGTSGGVLVCQTKHPWADGTSCEG 120
QY 538 KWCINGKCVNKHGKHFDTFPHGSGMWGPGWDCSRTCCTGGGVQVYTMRECDNPVPKNGGKY 597
Db 121 KWCINGKCVNKTDRKHFDTPFPHGSGMWGPGWDCSRTCCTGGGVQVYTMRECDNPVPKNGGKY 180
QY 598 CBGKRVYRSCNLEDCPDNNNGKTFREOCEAHNEFSKASFGSPGPAVEMIPKYAGVSPKDR 657
```

```
Db 181 CEGKRVYRSCNLEDCPDNNNGKTFREOCEAHNEFSKASFGSPGPAVEMIPKYAGVSPKDR 240
QY 658 CKLICQAKGIGYFVLQPKVVDGTPPCSPDSTSVCVQGCVKAGCDRIIDSKKFDPKCGVC 717
Db 241 CKLICQAKGIGYFVLQPKVVDGTPPCSPDSTSVCVQGCVKAGCDRIIDSKKFDPKCGVC 300
QY 718 GNGSTCKKISGTSVTSAPGVDHDIITPTGATNIEVKQNRGRNNGSFLAIKAADGTY 777
Db 301 GNGSTCKKISGTSVTSAPGVDHDIITPTGATNIEVKQNRGRNNGSFLAIKAADGTY 360
QY 778 ILNGDYTLSTLEQDIMYKGVVLRYSYGSSAALERIRSFPLKEPTIOVLTVGNALRPKIK 837
Db 361 ILNGDYTLSTLEQDIMYKGVVLRYSYGSSAALERIRSFPLKEPTIOVLTVGNALRPKIK 420
QY 838 YTYFVKKKESFNAIPTFSAWVIEEWGECSSKBLGQRRRLVECRDINGOPASECAKEV 897
Db 421 YTYFVKKKESFNAIPTFSAWVIEEWGECSSKBLGQRRRLVECRDINGOPASECAKEV 480
QY 898 PASTRPCADHPCPQWLQGEWSSCKTCGKGYKTSLSKLSHGGVLSHSDCDPLKPKPHF 957
Db 481 PASTRPCADHPCPQWLQGEWSSCKTCGKGYKRSLSKLSHGGVLSHSDCDPLKPKPHF 540
QY 958 IDFCTMAECS 967
Db 541 IDFCTMAECS 550

RESULT 3
T00355
hypothetical protein KIAA0688 - human
C:Species: Homo sapiens (man)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
C:Accession: T00355
R:Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.;
DNA Res. 5, 169-176, 1998
A:Title: Prediction of the coding sequences of unidentified human genes. X. The complete
A:Reference number: Z14142; MUID:98403880; PMID:9734811
A:Accession: T00355
A:Status: preliminary; translated from GE/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-837 <ISH>
A:Cross-references: UNIPROT:O75173; EMBL:AB014588; NID:g3327189; PIDN:BAA31663.1; PID:g3:
A:Experimental source: brain
C:Genetics:
A:Gene: KIAA0688
F:519-575/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match 40.3%; Score 2129; DB 2; Length 837;
Best Local Similarity 49.2%; Pred. No. 5.6e-135;
Matches 412; Conservative 132; Mismatches 224; Indels 70; Gaps 18;

QY 36 LLLIAALLAVSALGHPSEDEELVVPEL---ERVPGHGT-TRL---RLHAFDQQLDLV 89
Db 37 LLLILLASLLPSARLASPLPREBEIVPPEKLNGSVLPFGSGTPARLLCRLOAFGTLLLEL 96
QY 90 PPSSFLAPGFTLQNVGRK---SGSDTPLPLTDLAHCFYSGTVNGDPSSAAALSICEG- 144
Db 97 EQDSGVQVEGVTVOYLQGAPELLGGAE---PGT-----YLTGTINGDPESVASLHWGGA 148
QY 145 VRGAFYLLGDAYFTIQLPFAASERLATAAPGKPPAPLQFHLLRRNRQDVGGTGCVVDE 204
Db 149 LLGVLYRGAELHLQPLEGGTPNSA-GPGCA-----HILRRK----- 184
QY 205 PRPTGKAETDEDEBGTGEDEGEPQWSQDPALQGVQPTGTGSRKRFVSSHRYVETML 264
Db 185 -----SPASGGPGMNVKAP-----LGSPSPRPR-RAKRFASLSRFVETLV 224
QY 265 VADOSMAEPHSGGLKHVLLTLFSVAAALYKHPSIRNSVSLVWVKILVIHDEOKPEVTSN 324
Db 225 VADDKMAAFHAGLURVLLTVMAAAAKFHPHSIRNPVSLVTVTRVTLTGSGEBGQVGPS 284
QY 325 AALTLRNFNCWQKHNPPPSQSDAEHYDTAILFTRQDLCSQTCDTLGMADVGTVCDFSR 384
```

```
Db 285 AAQTLRSFCAMGRLTPEDSDPHFDFTAILFTRODLGCVSTCDTLGMADVGTVCDDPARS 344
Qy 385 CSVIEDDGLQAFTTAHELGHVFNPHDDAKOCASILING-VNODSHMMASMLNSLDHNSOPW 443
Db 345 CAIVEDDGLQSAFTAAHELGHVFNMLHNSKPCISLNGPLSTSRHVMAPVMAHVDPBEPW 404
Qy 444 SPCSGYMTSFLDNGHGECLMDKPNPIQLPGDLPGTSYDANRQCQFTFGEDSKHCPDAA 503
Db 405 SPCSARFITDFLNGYGHCLLKPEAPLHLPTVTFPGKYDADRQCQLTFGPDSPHCPQLP 464
Qy 504 STCSTLWCTGTSGGVLVCTQKHPHWDGTSCEGKWCINGKCNKVRKHKHFDPFHGSGW 563
Db 465 PPCAALWCSGHLNHAMCQTKHSPWADGTPCPAQAQCMGGRCLHMDQLDQDENIPQAGGNG 524
Qy 564 MNGPMDGCSRTCGGQVQYTMRECDNPVPKNQGGYCEGKRVYRNSNLEDCPDNNNGKTFRE 623
Db 525 PNGPMDGCSRTCGGQVQYFSSRDTCTVPKNGGKYCEGRTRFRSNTEDCTPGSALTFR 584
Qy 624 EQCEAHNEFSKASFGSPA-VEMI PKYAGVSPKDRCKLICQAKGIGYFVLPQKVVVDGTP 682
Db 585 EQCAAYNHRDIL-FKSFPGPMDVPRYTGVAPQDQCKLTCQARALGYVYVLEPRVVDGTP 643
Qy 683 CSPDSTSVQCQCVKAGCDRIIDSKKPKDKCGVCGGNGSTCKISGVSFSAPKGYHDII 742
Db 644 CSPDSSSVQCQRCITHAGCDRIIGSKKPKDKCMVCGGSGGSGKSGSFRKFRYGYNNV 703
Qy 743 TIPTGATNIEVKNORNGSRNNGSFLAIIKADGTYLNGDYTLSTLEODIMYKGVV-LRY 801
Db 704 TIPTGATHILVROQNGPHGRS--IYALAKLPDGSVALNGEYTLMPSPDVLPGAVSURY 761
Qy 802 SGSSAALERSFSPLKEPLTIQVLTIVGNALRPKIYTYFVKKKKESFNAIPTFSAW 859
Db 762 SGATAASETLGSHGLPLAQPLTIQVLVAGNPQDTRLRYSPFVRPTPS-TPRPTPDWL 818
;
RESULT 4
T21371
hypotheetical protein F25H8.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T21371; T24896
R:Gajadaty, S.
submitted to the EMBL Data Library, February 1996
A:Reference number: Z19413
A:Accession: T21371
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2165 <WIL>
A:Cross-references: UNIPROT:Q19791; EMBL:Z69360; PIDN:CAA93287.1; GSPDB:GN00022; CESP:F2
A:Experimental source: clone F25H8
R:Gajadaty, S.
submitted to the EMBL Data Library, February 1996
A:Reference number: Z19949
A:Accession: T24896
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2165 <W12>
A:Cross-references: EMBL:Z69361; PIDN:CAA93288.1; GSPDB:GN00022; CESP:F25H8.3
A:Experimental source: clone T13H10
C:Genetics:
A:Gene: CESP:F25H8.3
A:Map position: 4
A:Introns: 31/1; 52/1; 135/2; 193/3; 216/1; 266/1; 495/2; 547/3; 584/3; 634/2; 744/1; 81
Query Match 30.2%; Score 1598; DB 2; Length 2165;
Best Local Similarity 32.0%; Pred. No. 8.4e-99;
Matches 355; Conservative 164; Mismatches 355; Indels 236; Gaps 29;
Qy 48 DALGRPSEE-----DEELVPELERVPVGHGT-TRLR-----LHAFDQQLDLDVDPD 92
Db 78 DSYGRGRKRKRDVTSTDRRRRLQGVARDGCHLRRLRSDDAVYIVLHRWNQ-----IPDS 132
```

```
Qy 93 SSFLAPGFTLQNGR-----KSGSDTPIPETDLAHCYFGTGVNGDPSSAAALSL 141
Db 133 HNKSVPHSNNGFAPMWLYLDSEBEVRGMSRTDP-----CIYRAHVKG-VHQHSIVNL 186
Qy 142 C---EGVRGAYLLGEAYFIQPLPAASERLATAAAGEKPPAPLOPHLLRRNRQDVGTC 198
Db 187 CDSBDGLYGMALPSGIHTVEPIISGNGTEHDGASRHRQ-----HLVRK----- 230
Qy 199 GVVDDEPRPTKABTEDEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGE 257
Db 231 ---FDPMHFKSFDHLNSTSVNETETTVATWQDQ-----WEDVIERKASRRASNSWD 279
Qy 258 RYVETMLVADOSMAEFHSGSLKHYLLTLFSVAARLYKHPISRNSVSLVVKILVHDBQK 317
Db 280 HYVEVLVVADTKMYEYHRSLESDYVTLFSTVASIYRHSQSLRASINVVVVKLILVTENA 339
Qy 318 GPEVTSNAALTRNFCWNQKOHNPSPDRDAEHYDTAILFTRODLGSGQ--TCDTLGMADV 376
Db 340 GPRITQNAQTLQDFCRWQYINPDSSVQHHDAVALLTRKDCRSQGKCDTLGLAELG 399
Qy 377 TVCDPSRSCSVIEDDGLQAFTTAHELGHVFNPHDDAKQACASLNGVQ----- 425
Db 400 TMDWQKSCAIIEDNGLSAAFTIAHELGHVFSIPHDDERKCSYMPVNVKCKFQSTKFDK 459
Qy 426 -----DSHMMASMLNSLDHNSQSPSCSGYMTSFLDNGHG--ECLMDKPNQPIQLP 475
Db 460 TQFNQNFHIMATPLEYNTHPMSWSPCSAGMLERFLENRGTQCLFDQFVRRRYEDV 519
Qy 476 -DLQCTSVDANROQFTFGEDSKHCPDAASTCSTLWCTGTSGGVLVCTQKHPWADGTSC 534
Db 520 RDEPGKYDAHQCKFVFGPASELCP-YMPTCRRLWCATFYGSGQMGCKTQIMPWADGTPC 578
Qy 535 GEGK--WCINGKCNKVRKHKHFDPFHGSGWGMWGPWGDSCRTCGGQVQYTMRECDNPVK 592
Db 579 DESRMFCHGACVRLAPESL--TKIDQWGDMSWGECSTCGGQVQKGLRDCDSCP 636
Qy 593 NGKYCEGKRVYRSCNLEDCPDNNNGKTFREEQCEAHNEFSKASFG-----SGPAVEWIPK 648
Db 637 NGKYCVQORRYRSCNTQCEPMDT-QPYREVQC---SEFNKDIGIQGVASTNTHWVPK 692
Qy 649 YAGVSPKDRCKLICQAKGIGYFVLPQKVVDPGTPCSPDSTSVQGVQCVKAGCDRIIDSK 708
Db 693 YANVAPNERCKLYCRLSGSAAPYLLRDKVVDGTCDRNGDDICVAGACMPAGCDHQLHST 752
Qy 709 KFKDKCGVCGGNGSTCKKISGVSFSAPK-GYHDIITITPTGATNIEVKNORNGSRNNGS 767
Db 753 LRRDKCGVCGGDDSCCKVVKGTFTNEQGTFGYNEVMKIPAGSANIDIRKGYNNMKEDDNY 812
Qy 768 LAIKAADGTYLNGDYTLSTLEQDIMYKGVVLRYSGSSAALERIRSFPLKEPLTIQVLT 827
Db 813 LSLRAANGEFLNGHFQVSLARQQTAFQDVTLEYSGSDAIIERINGTGPISRSDIYVHVLS 872
Qy 828 VGNALRPKIYKYTF----- 841
Db 873 VGSH-PPDISYEYMTAAVPNAVIRPISALYLWRVTDWTWTECDRACRQOQKMLCLDMS 931
Qy 842 -----VKKKESFNA--IPTFSAMVIBEWGECSSKSCELGMORRLVECRDINGQ 887
Db 932 THROSHDRNCNVLPKPKQATRCMNCIDCSTWITEDVSSCSAKCGSGQKQRQVSCVMESG 991
Qy 888 ---PASE----- 891
Db 992 RQTPASEHLCDRNSKPSDIASIDCSGRKNYGEWTSCTSETCGSNGKWRHRSKVCVDDSN 1051
Qy 892 -----CAKEVKKASTPCADHPCPQWQLGEMWSSCKTKCGKYKKTSLKCLSHDGGVLS 944
Db 1052 RVDSESLCGRGOKATERECNRIPCPRWVYGHWSECSRSCTGGVGMRRHAQCLDAADRETH 1111
Qy 945 HDSCDPLKKPKH-----FIDFCTMAECS 967
Db 1112 TSCGPAQTQEHCHNEHACTWQFGVWSDCS 1141
```


Db 397 ETVDCEAEWFTGDMWSSCSCTGCGQYRVVYCHQVFANGRRVTVEDGNCTVE-RPPVKQ 455
 QY 903 PCADHPCPQWQGEWSSCSKTCGKYKTTSLKCLS-----HDGVLSHSDCDPLKPKPHFI 958
 Db 456 TCNRFACPEWQAGPWSACSEKCGDAFQYRSVTCRSEKEGEGKLLAADACPADEQEKFT 515
 QY 959 D-FCMAEC 966
 Db 516 ERTCNLGPC 524
 RESULT 7
 T34395
 hypothetical protein C37C3.6b - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T34395, T34394
 R:Geisels, C.; Bradshaw, H.
 submitted to the EMBL Data Library, July 1996
 A:Description: The sequence of C. elegans cosmid C37C3.
 A:Reference number: Z21518
 A:Accession: T34395
 A:Status: preliminary; translated from GB/EMBL/DBBJ
 A:Molecule type: DNA
 A:Residues: 1-2167 <GEI>
 A:Cross-references: UNIPROT:O76840; EMBL:U64857; PIDN:AAC25868.1; GSPDB:GN00023; CESP:C3
 A:Experimental source: strain Bristol N2; clone C37C3
 A:Accession: T34394
 A:Status: preliminary; translated from GB/EMBL/DBBJ
 A:Molecule type: DNA
 A:Residues: 1-1555, 'SKP' <GE2>
 A:Cross-references: EMBL:U64857; PIDN:AAC25867.1; GSPDB:GN00023; CESP:C37C3.6a
 A:Experimental source: strain Bristol N2; clone C37C3
 C:Genetics:
 A:Gene: CESP:C37C3.6b; CESP:C37C3.6a
 A:Map position: 5
 A:Introns: 32/3; 104/2; 156/2; 207/3; 459/2; 536/3; 577/2; 1105/3; 1367/1; 1438/1; 1556/

Query Match 12.2%; Score 645; DB 2; Length 2167;
 Best Local Similarity 31.3%; Pred. No. 8.3e-35;
 Matches 153; Conservative 73; Mismatches 163; Indels 100; Gaps 18;
 QY 546 VKNHRKHFDTPFHSGWGMWPGDCSRTCCGGVQVYTMRECDNPVPMKNGKCYCEGRVRY 605
 Db 68 VDKSGQKE-----TGNWGPWPVENECSRCGGVQLEKRCQSGD-----CTGASVRY 114
 QY 606 RSCNLEDPCDNGKTFREOQCAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLICQAK 665
 Db 115 ISCNLNAE--ESGTFDPRABQCSKFNDAL-----DGNVHKWTP-YKG---KNKCELVCCKPE 164
 QY 666 GIGYFVLQPKVVDGTFPCSPDSTSVCQOCQKVGACDRIIDSKKPKDCKGCGNGSTCK 725
 Db 165 SGNFYKWKADKVDGTCDSKNDICVDGECPLVPGCDGLGSLKFDCKGCKDGGSTCK 224
 QY 726 KISG--SVTSAPKGYHDIIITPGATNIEVKQNRGSRNNSFLAIKAADGTYILNGDY 783
 Db 225 TIEGRFDERNLSPGYHDIIKLPAGATNIQES--ARKSTNN---LALKNGSDHFFVLNGN- 278
 QY 784 TLSTLEQDIWYKGVVLRYSGSSAALERIRSPLKEPLTIQVLTGNALRPKIKTYFYVK 843
 Db 279 GLIQVEKEVEGGTIPVY--DDAEPETLSAQPLSELTVALLFKRGSRDTAIFYEFSIP 336
 QY 844 KKKE-----
 Db 337 LBEEDVDMYKFDNWTPCSVSGKGVQTRNLICIDGNKGRVEDDLCEENNATKPFPEKSC 396
 QY 850 NAIPTFSAWVIEWEGECSKSC-ELGQRRRLVBCRDI--NGQPAS-----EKAKEVFPASTR 902
 Db 397 ETVDCEAEWFTGDMWSSCSCTGCGQYRVVYCHQVFANGRRVTVEDGNCTVE-RPPVKQ 455
 QY 903 PCADHPCPQWQGEWSSCSKTCGKYKTTSLKCLS-----HDGVLSHSDCDPLKPKPHFI 958

Db 456 TCNRFACPEWQAGPWSACSEKCGDAFQYRSVTCRSEKEGEGKLLAADACPADEQEKFT 515
 QY 959 D-FCMAEC 966
 Db 516 ERTCNLGPC 524
 RESULT 8
 T18856
 angiogenesis inhibitor homolog - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T18856; T24653
 R:McMurray, A.
 submitted to the EMBL Data Library, July 1995
 A:Reference number: Z19031
 A:Accession: T18856
 A:Status: preliminary; translated from GB/EMBL/DBBJ
 A:Molecule type: DNA
 A:Residues: 1-1444 <WIL>
 A:Cross-references: UNIPROT:Q8MYA8; EMBL:Z50004; PIDN:CAA90293.1; GSPDB:GN00028; CESP:C0
 A:Experimental source: clone C02B4
 R:McMurray, A.
 submitted to the EMBL Data Library, July 1995
 A:Reference number: Z19917
 A:Accession: T24653
 A:Status: preliminary; translated from GB/EMBL/DBBJ
 A:Molecule type: DNA
 A:Residues: 1-1444 <W12>
 A:Cross-references: EMBL:Z50006; PIDN:CAA90302.1; GSPDB:GN00028; CESP:C02B4.1
 A:Experimental source: clone T07C5
 C:Genetics:
 A:Gene: CESP:C02B4.1
 A:Map position: X
 A:Introns: 25/3; 70/3; 96/3; 139/3; 187/1; 234/2; 282/3; 376/2; 422/2; 478/3; 509/3; 566/

Query Match 11.6%; Score 614.5; DB 2; Length 1444;
 Best Local Similarity 23.8%; Pred. No. 5.6e-33;
 Matches 237; Conservative 121; Mismatches 336; Indels 291; Gaps 45;
 QY 62 VPELVRVPG-----HGTVRLRLHAFQOQLDLDVPPDSSFLAPGFTLQNVGRKSGSD--T 113
 Db 26 VPEVSLIEATRHPLKNGLNKMKFTAWNDYHLNLRKRSRVSP--HIISVVVRHGDDVTT 83
 QY 114 PLPETDLAHCFYSGTVNGDPSSAALSCEGVGAFYLLGAYFIQPLPAASERLATAAP 173
 Db 84 YAGLRDYEQCHYQGEVSKSHGNKAAISDCGALMSIVMEDHFLVLQTLPK----- 133
 QY 174 GEKPPAPLQFHLLRRNRQDVGTCGVVDDEPRPTGKAETEDEGTEGEDEGPQWSP-- 231
 Db 134 -----RVHHLQKERHLVYKRSAGLLTN-----AESKIRBEITRLQEEQESFCDTS 178
 QY 232 --QDPALQ-----GVGQPTGTG-----SIRKKRFVSS-----HRYVETML 264
 Db 179 EQLDDPANTIPAHLHFNVTIPTSALQDSSFTFPNMDPITLIGLFLDSKLFHEHREYIQ 238
 QY 265 VADQSMAEFHSGGLKHYLLTLFSVAARLYKHPISRNVSLSVYVVKILVTHDE-----QKQPE 320
 Db 239 DAEGHLLFEFS-----LALINNHHVLYQDQTLTPNLDIVIVRYEMWRTQPSALSSTGVH 290
 QY 321 VTSNALTLRNFQWQKQHNPPSD-RDAEHVDYDTALLFRQDLCSGQTCDDTLGMADVGTVC 379
 Db 291 KNGQAQSLDFAFCRYQAHMNFQTDLTDMNHVDHGVLLTGYDIYHT-TTTSVAGVAPVARMC 349
 QY 380 DPSRSCSVIEDGQAAFTTAHELGHVFNMPHDDAK-OC-----ASLNGVNDQSHMA 431
 Db 350 DPLFACSLVEGLHGRSFVLAHEMHNMGVHDGVQNCNKGCCCLMSAVNGAKTT----- 405
 QY 432 SMLSNDLSQPSWSPSCGYMITSFL-----DNHGCECLMDKPNQPIQLPG-----DL 477
 Db 406 -----WSDCSVREFNAFLQLQDESGRGNCLRDAS-----PGLISTNHLSLDRL 448
 QY 478 PGTSYDANRQOFTFGEDSK-HCPDAA-----TCSTLWCTGTSGGVLCQTKHFPWADGTS 533

Db 449 PGORFTADQCSYFNGRDYKVEIPNGKAMDDICRIILWC-GNSGSTI--STAHPALEGSW 504
QY 534 CSEGKWCINGKCVNKNRKHFP-----DTPFHGSWGWG-----PWGDCSRCTCGGVQVT 582
Db 505 CGANKWCHKGOCT-----HWTFGLTTPVPIDGSEWGAEGKCPICQCAVSGSITVQOQ 558
QY 583 MRECDNPVKNKGKCEGKRVRYSNLEDCPDNNGKTFREBQCEAHNEFSKASFGSGPA 642
Db 559 HRDCVNPAPNNGKTCGEGNIRGIVC-----GAT--SSNCLG--FTRFEEFGN--- 601
QY 643 VEWIPKIYAGVSPKDRCKLIQAAGIGYFFVLQF-----KVVDGTPGSP 685
Db 602 -----KICSSIKYDPHKPDQQLTGEGFHSHTQPCRWCHLIGSELLRNKGQFPDGTGPGF 656
QY 686 DTSVCVQOQCVKAGCDR--IIDSKKFKDKCGVCGNGSTCKKISGVSPTSAPGVH--DI 741
Db 657 D--AYCVGGQCLALSCDNKALVEQPE-----DCPRIEGRSVHOWEESWSE 701
QY 742 ITIPTGATNIEVKORN-ORGSRRNGSFLAIAKAADGTIYLNGDYTLSTLEQDIMYKGVVLR 800
Db 702 CSVSCGLGREVREKCSGRKCGQVSESRPCEG-----VLR 739
QY 801 YGSSAALERIRSFPLKEPLTIQVLTGNALRPKIKYTYFYKKKESFNAIPTFSAWVI 860
Db 740 -----DCEERGEW-- 747
QY 861 BEWGECSKSELGWORRLVECRDINGQPASECAKEVKPASTPCADHPC-PQWQLGEWSS 919
Db 748 KEWGSCEKCALGVQKFRPC--LTDQCSSHLQE-----BRPCDNEGCWTNWD--EWSS 798
QY 920 CSKTGCGKGYKTSKLSLASH--DGGVLSHSDCDPLK 952
Db 799 CSQSCGGRRYIRKCLDDKCDGDDLEXESCNTQK 833

RESULT 9

T16892
hypothetical protein T19D2.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T16892
R:Bentley, D.
submitted to the EMBL Data Library, December 1995
A:Description: The sequence of C. elegans cosmid T19D2.
A:Reference number: Z18599
A:Accession: T16892
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-860 <BEN>
A:Cross-references: UNIPROT:Q22580; EMBL:U42846; NID:g1125810; PIDN:AAA836
C:Genetics:
A:Gene: CESP:T19D2.1
A:Introns: 56/3; 96/1; 171/3; 199/3; 245/2; 293/2; 335/3; 395/3; 426/3; 484/2; 505/3; 58

Query Match 8.8%; Score 466; DB 2; Length 860;
Best Local Similarity 23.5%; Pred. No. 2.7e-23;
Matches 180; Conservative 100; Mismatches 319; Indels 168; Gaps 37;
QY 248 IRKRFVSSHR-----YVETMLVADQSM-----AEPHSGSLKHYLLTLFSVAAR-----LYKH 295
Db 119 VKDIRRAEPHRRDILTVELAFADDAMWDHFKOMYKAAEENMHTFIMAVVNNIDVLVYQ 178
QY 296 PSIRNSVSLVVKVLVI-----HDEQKGPEVTSNAALTIRNFCNWKQKHNPSPDRDAEYD 351
Db 179 RLLQPRINIKIVRYEILKNIIPHLNARKHSGDVRDLDAFCQYQNEINPNPDADPRWD 238
QY 352 TAILFTRODLCSOTCDTLGMAADVTCVDPKRSKSVIEDDDGLQAAFTTAHELGHVFNWPH 411
Db 239 HALLFSGYDLHRNGVKTVAGYAPVKMCSGVRSCTINEGLDFGVVVVTHMGHSLGMYH 298
QY 412 DDAKQCASLNGWQDSH-----MMAASMLNLDHQPWSPGCSGYMITSFLDNHG-----GE 461

Db 299 DGDNEC-----DLRCCIMSPSVSGKTH---WSQCSVNMATFV--GHLGDFFRPPN 345
QY 462 CLMDKPQNPQIOL-----PGDLPGTSDANRQCOFTFGEDSKHCPDAAST-----CSTLWCTG 513
Db 346 CLQDASANEQRMVAFKESPEPQLFTLDBQCEIIPHGECKWHELKDGQTMQMCQWMC-G 404
QY 514 TSGGVLCVOTKHFHWADGTCGEGKWCINGKCVNKNRKHFPDTPFHSGSWGMPGWD-CS 572
Db 405 NGEVVI--RTAH-PALEGTTCGFGMI CRQGCVCSSQLMRVT-----GGWSTWDRPA 455
QY 573 RTCGG-----GVQYTMRECDNPVKNKGKCEGKRVRYSNLEDCPDNNGKTF 621
Db 456 PTCGRCSCQCEIRQIRIMRSIROCNPNSSNNGGAPCQGDGEARGMVCHRDVCNGDSIENY 515
QY 622 REEQCEAHNEFSKA--SFGSGPAVEWIPKYGAVSPKDRCKLIQAAGIGYFFVLQPKVVD 679
Db 516 ATRVCSRLRDENAIPTILSSEGMOF-----EAMCKTIWCLISGTSNIRTYS-NFPD 566
QY 680 GTPCSPDSTSVCOGQCVKAGCDRIIDSKKFKDKCGVCGNGST-----CKKISG 729
Db 567 GAPCGPG--QYCIKGECPILC-----GSTLAYSEADCPVLQVOT 605
QY 730 SVTSAKPGYHDIITPTGATNIEVKORNQGRSNGSFLAIAKAADGTIYLNGDYTLSTLE 789
Db 606 TTTTPPMHVHSVDQF-AGKTN-PYKEHKTPFLNWSGWSVMSRCVTY----- 651
QY 790 QDIMYKGVVLYRSSSAAAL--ERIRSFPLKEPLT--IQVLTVG--NALRPKI----- 836
Db 652 -DCHTQGVKVRVRRCLAGVCAGALRERQPCTRPCTGSRPLTTSPPOOTFRNRFIAPLPN 710
QY 837 KYTVFVKKESFNAIPTFSAWVIEEWGECSSKSELGWQ--RRLVECRDINGQPASECAKE 895
Db 711 RQTNMILRKVDHWG-----PMSACSVTCGTGOKLRRENC---TGO---ECAA- 752
QY 896 VKPASTRPCADHPCPQ-----W-QLGEWSSCSKTCGKYKTSKLSKLS 937
Db 753 -----TGFCVMQSCRENKNTWTEWSQWSDCVNCGEGVQFRKACFA 794

RESULT 10

T22545
hypothetical protein F53B6.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T22545
R:White, S.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19578
A:Accession: T22545
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1059 <WIL>
A:Cross-references: UNIPROT:P90884; EMBL:Z81086; PIDN:CA803121.1; GSPDB:GN00019; CESP:F5;
A:Experimental source: clone F53B6
C:Genetics:
A:Gene: CESP:F53B6.2
A:Map position: 1
A:Introns: 38/3; 92/3; 131/3; 169/1; 236/2; 282/3; 349/2; 420/1; 576/3; 759/3; 84;

Query Match 7.9%; Score 418.5; DB 2; Length 1059;
Best Local Similarity 24.6%; Pred. No. 5.5e-20;
Matches 102; Conservative 66; Mismatches 121; Indels 125; Gaps 16;
QY 561 SWGMWGPWGDCSTCGGVQYTWRECDNPVKNKGKCEGKRVRYSNLEDCPDNNGKT 620
Db 25 SWAAWSPWSSCTTKTCGGVSRQLRRCLT-----SKCSGESVRFKVC----- 65
QY 621 FREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLIQAAGIGYFFVLQPKVVDG 680
Db 66 -AQKTCESKSLARDITCGGEEI-----VS-RGQCEVVCRSRLTGANFLW--RVDDG 113
QY 681 TPC-SPDSTSVCOGQCVKAGCDRIIDSKKFKDKCGVCGNGSTCKKISGVSPTSAPGVH 739

Df 114 TPCQAATSRAVCSKSGCQIVGCDGLISSRFFDAGCVCGGRGDTC----- 158

Qy 740 DIITPTGATNIEVKORNORSGRNNGSFLAIKAADGTGYILNGDYTLSTLEODIMYKGVL 799
||| : : : : :
Db 159 -----DNGKFI-----WKVEEYTACASNC-----DIV 182

Qy 800 RYSGSSAALERIRSFSLPELTIQVTV-----GNALRPKIYTFVKKKESF 849
||| : : : : :
Db 183 DWSGAG-----RSIASTSQPIVVVCNAITGRVVPEKLCADKLRPKVE-----A 225

Qy 850 NAIPTF---SAWVIBEWGECSKCELGWQRRLVECRDINGO-----PASECAKEVKPAST 901
||| : : : : :
Db 226 RPCPMLICSRMAAWTECPHCIGEGTRREVYCQTAAHNVTVHVDPDFCEGTRPAAE 285

Qy 902 RPCADHPCCQWLGEWSSCKTGGKYKTSLKCLSHDGGVLSHSDCDPLKKPK 955
||| : : : : :
Db 286 ENCUSTSCGRWEAGKWSKCTASCQGVRRRHVACV--GG----SDCDEGGRPR 332

RESULT 11

Tl5976
hypothetical protein F08C6.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: Tl5976

R;Bentley, D.
submitted to the EMBL Data Library, June 1995
A:Description: The sequence of C. elegans cosmid F08C6.
A:Reference number: Z18440
A:Accession: Tl5976
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-957 <BEN>
A:Cross-references: UNIPROT:Q19204; EMBL:U29378; NID:g868184; PID:g868185; PIDN:AAA68721
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:F08C6.1
A:Introns: 23/1; 135/3; 220/3; 285/3; 325/2; 376/3; 461/2; 498/3; 532/3; 627/2; 714/1; 714/1;

Query Match 7.6%; Score 402; DB 2; Length 957;
. Best Local Similarity 21.3%; Pred. No. 6.2e-19;
Matches 227; Conservative 122; Mismatches 384; Indels 354; Gaps 53;

Qy 78 LHAFQQLDLDPD--SFLAPGFTLNVGKSGSDTLPET-DLAHFYSGTGVNGDPSS 135
||| : : : : :
Db 61 IQAFNKYNLSLEXTLAKLLSSGVTVVKNKKGSLDFGSTLDSCHVHHYG-----EKV 115

Qy 136 AAALSCEG-----VRGAFYLGEAYFTOLP---AASERLATAAPEKPPAPLQ 182
||| : : : : :
Db 116 YAAISNCGRIVESNRKQKGTVIDDGEIIVVHPFDHHAHRKRATENG----- 167

Qy 183 FHLL--RRNRQDVGCCTGVVDDEPRTPGKAETD---EDBGTEGDBGQWPSPDPALQ 237
||| : : : : :
Db 168 -HVVYRETLAGEKDFCGLI-----DNVTTESLIVEDESAIFEV----- 206

Qy 238 GVGPQTGTSIRKKRFVSRR-----YVETMLVDAQSM-----AEFHGS---GLKHYL 282
||| : : : : :
Db 207 -----FVTGORLTQQSLDI VELA VFVDENLWRHFSSKGGGMADRKLQDYT 251

Qy 283 LTLFSVAARLYKHPSIRNSVSVLVVVKLVTHDEOKGPEVTS-----NAALTIRNFCNW 335
||| : : : : :
Db 252 LTLLNNIQIMYQTPASPPLTFRIVREVL---TRQPSALAGYLHNHGNAQMAYLDRFCRY 308

Qy 336 QKQHNPSPDRDAEHYDTAILFTQDL---CGSQTCIDLGMADVCTCDPSRSCSVIEDDG 392
||| : : : : :
Db 309 QRN---LAVRDWDH---AIMLTGYDIRHAGSRISIS--GIARLDGMCDPWMTCTLAEGLD 360
||| : : : : :

Qy 393 LQAAPFAHELGHVFNNPHDDAK-----QCASLNGVNO-----DS 427
||| : : : : :
Db 361 FTSAFIGTHELGHFRFTELKDHKSDTLTGTCFGPSKWCOLRCVPWTGTNELIOPTVQVHA 420
||| : : : : :

Qy 428 HMMASMLSNDHSQPWS-----PCSGYMITSFLDNGHBCECLMDKPQNPIQLGBDL 477
||| : : : : :

421 PVVTTLPSRIDGS--WSGWGATICSQCTNGTILSGVLAIARTCSAPYPANG---GSDC 475

478 PGTSYDA---NRQOFTFEGEDSK-----HCPDAAASTCSTLMTGTGSGVLVCQTKHF 526

476 VGSTSRAVLCSRQC---GRASKSVDEYISDKCMQKRLKNDRELTKGS-----QLNRF 526

527 P-----WADGTSCEGKWKINGKCVNKHKHK----- 553

527 PQRACKVFCVQOHYGSQRNYRFFGDNLPDGTSCGYDRYCLDGECLALNCNNALISRDQ 586

554 -----FDTPFHSGMGPWGDCSRTCGGGVQYTMRECDNPPVKGNGKVCCEGK 601

587 SCPTDTCPIITQSSSVYKQGTWLSLTCTATCGGYRKRNRACISIT-----GQCEGN 640

602 RVYRSCNLEDCPDNNNGTKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKDRCKLI 661

641 EDETEVCSSSESCP-----SVLRVGNWSTWT-----EW-----NHCSVS 674

662 C-----QAKGIGYFFVLQPKVVDGTPCSPDS-----TSVCVQGCVKAGCDR 703

675 CGRGSQAR---YRKCLSPHRTLAFDCPENKVTNELRIITFFKARSYIMCSVRCNKIKRNT 731

704 IIDSKKKFKDKC-----GVGC--GNGSTCKKISGTSVTSAPKPGYHDIITPTGATNIEV 753

732 ISEKNIEVRSCDGPACNAIGVGTGWGHWSTCS-----TSCGPG-----TL-----V 772

754 KQRNQRGRNNGSF-----LAIKAADGTYLNGDYTLTLEQDIMYKGVVLRYSGSSA 806

773 RQRTCNREPCDGAHERRSCNVATCQNDGIWSLWNEWSDCS---RVCGKGL----- 820

807 ALEIRISFPLKEPLTIQVLTVGNALRPKIKYTVFKKKKESFNAIPTTSAMWIEWGEC 866

821 ---RSRSRS-----CFGSGCMGASSRQPCNEQACASSANDMGWTW--SGWSQC 864

867 SKSCELGNQRELVBECRDINGQPASECAKEVPASTRPCADHPCPQ-----WQLGEWSS 919

865 SVSCGAGVKRRTRTCRTGN-----CPGNYKESAI--CNDRDCENKNAAWGGW--GIWSS 914

920 CSKTCGKGYYKKTSLKLKSHDGVLSHSDCLPKLKKPKHFIDFCTMAEC 966

915 CSETCGDGVRRVRKC-----YSGNGCDGOQYEK---QYCNLRVC 951

RESULT 12

T00260

hypothetical protein KIAA0605 - human

C;Species: Homo sapiens (man)

C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004

C;Accession: T00260

R;Nagase, T.; Ishikawa, K.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.; Ohara, O. DNA Res. 5, 31-39, 1998

A;Title: Prediction of the coding sequences of unidentified human genes. IX. The complete

A;Reference number: Z14086; MUID:98290545; PMID:9628581

A;Accession: T00260

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-951 <NAG>

A;Cross-references: UNIPROT:O60345; EMBL:AB011177; NID:G3043733; PIDN:BAA25531.1; PID:G3043733

A;Experimental source: brain

C;Genetics:

A;Note: KIAA0605

F;46-106/Domain: thrombospondin type 1 repeat homology <THRI>

Query Match 7.4%; Score 391; DB 2; Length 951;

Best Local Similarity 22.8%; Pred. No. 3.4e-18;

Matches 134; Conservative 61; Mismatches 194; Indels 200; Gaps 21;

Qy 562 WGMGMPGDCSRTCGGGVQYTMREC-----DNPVPKNGGKYCEGKRVYRSCNLEDCPDNN 617

Db 50 WGEWTKTAFSRS CGGGVTSQERHCLQRRKSVPGPNRTCTGTSKRYQLCRVQECPP-D 108

Qy 618 GKTFRSEQCEAHNEFSKASFGSGPAVEWIP-----KYAGVSPKDRCKLIICQAKGIGYFFVL 673

```
Db 109 GRSFEEQCVSFN-----SHVYNGRTHQWKPLYPDDYVHISSKP-CDLHCTTVD-QORQLM 162
Qy 674 QPKVVDGTPCS-PDSTSVQVQOCVKAGCDRIIDSKKFKDFKCGVCGNGSTCKKISGSVT 732
Db 163 VP-ARDGYSCKLTDLRGVCVSKCFPIGCDGLFSTHLDKGCICQGDGSSCTHTGNRY 221
Qy 733 --SAKPGYHDIITPTGATNIEVKORNGSRNNGSFLAIAAADCTYILNGDYTLSTLEQ 790
Db 222 KGNALHGLSVLTHIPAGARDIQUVER-----KKSADVLALADEAGYFFNGNYKVDSPK 275
Qy 791 DIMYKGVVLRYSGS-----SAALERIRSFSPKLEPTIOVLTVGNALRPKIKYTPVKKKK 846
Db 276 NFNIAGTVVKKRRPMDVYETGIEYVAOQPTNOGLNVMVN-ONGKSPSITFEYTLQPP 334
Qy 847 ESFNAIPFSAWIBE-----WGECSKSCELGHQRLV----- 879
Db 335 HESRPQPIYYGFSESAESQGLDAGIAGMFIPHNGSLYQOAS-SERLGLDNRLFGHPGLDM 393
Qy 880 -----EC-----RDIN----- 885
Db 394 ELGFSQOQETNEVCQAGGACGEPGRKGFDRDNRVTGTLTGDKDDEVDTHFASQEFF 453
Qy 886 -----GQPASECAK----- 894
Db 454 SANAIISDQLLAGSLDKDFTLNETVNSIFAQAPRSSLAESFFVDYBENEGAGPYLLNGS 513
Qy 895 -----EVKPASTRPCAD-----HPCP--OMQLGEW 917
Db 514 YLELSSDRVANSSBAPPFNVSTLSLTGAGNRTHKARTRPKARKOGVSPADMYRWKLSSH 573
Qy 918 SSCSKTCKGKYKTKSLKCLSHDGHGVLSDCDPLKKPKHFIDFCFMAEC 966
Db 574 EPCSATCTGWSAVAMCVRGDVEVDSDSYCDALTRPEPVHEFCAGREC 622
.
RESULT 13
S60257
meltin alpha - mouse
A:Species: Mus musculus (house mouse)
C:Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 09-Jul-2004
C:Accession: S60257
R:Yagami-Hiromasa, T.; Sato, T.; Kurisaki, T.; Kamiyo, K.; Nabeshima, Y.I.; Fujisawa-Seth
Nature 377, 652-656, 1995
A:Title: A metalloprotease-disintegrin participating in myoblast fusion.
A:Reference number: S60257; MUID:96026308; PMID:7566181
A:Accession: S60257
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-903 <YAG>
C:Cross-references: UNIPROT:O61824; EMBL:D50411; NID:G1054586; PIDN:BAA08912.1; PID:G105
C:Superfamily: mouse meltin alpha; disintegrin homology
F:421-503/Domain: disintegrin homology <Dis>
F:349/Active site: Glu #status predicted

Query Match 7.2%; Score 378.5; DB 2; Length 903;
Best Local Similarity 22.0%; Pred. No. 2.2e-17;
Matches 189; Conservative 102; Mismatches 278; Indels 289; Gaps 37;

Qy 21 AERAPGSRFGVPFTLLALLAALVSDALG-----RPSDEELVPELER 67
Db 6 ARRAPPARAL-----LLALAGALLAPRAARGMSLWDQRGAYEVARASILLSKDPGIPG-QS 59
Qy 68 VPGHG-----TTRLRLHAFDQOLDLVPDSSFAPGFT-----LQNVGRKSGSDTLPET 118
Db 60 IPAKDHPDLVTQVQLSERD--LILSLRNEGLIANGFTETHYLDQ-----GTDVSLTRN 112
Qy 119 DLAHCFYSGTVNGDPSSAAALSLCEGVRGAFYLLGEAYFIQPLPAASERLATAAPGEKPP 178
Db 113 HTDHCYHYGHVQGDAAASVVSLSTCSDLRGLMFENKTYSLRPMKNTD----- 160
Qy 179 APLOPHLLRRNRQGDVGTCGVDDDEPRPTGKAETDEDEGEDEGPQWSPQDPALQG 238
Db 161 ---SYKLVPAESMTNIOGLCG-----SQHNKSLNTMEDVSP----- 193
```

```
Qy 239 VGQPTGTGSIR-----KKRFVSSSHRYVETMLVADQSMAEFHGSG-----LKHVLLTLFSPA 289
Db 194 -----GTSQMRARRKHRETLMKTKVELVIVADN--REFQRGKDLEKVKQORLIEIANHV 246
Qy 290 ARLYKHPSIRNSVSLVVKILVIVHDEQKGPVETVNAALTLRNFNCWQKQHPSPDRDAEH 349
Db 247 DKFYRPLNIR-----IVLVGVEVWMDIK-CSISQDPPTFLRHEFLDWRKIKLLP-----RKS 297
Qy 350 YDTAILFTRODLCSQTCDTLGMADVGVTCVDPSPSCSVI-----EDDGLQAAATTAEHLGHV 406
Db 298 HDNAQLISGVYFQGT-----TIGMAPIMSMCTAEQSGGVMDHSDPSLGAAYTLAEHLGN 353
Qy 407 FNMPHDDAKOCASLNGVNDQSHMMASMLNLDHSQSPSCGYMITSFLDNGHGCECLMDK 466
Db 354 FGMNHDILERCSCSRMAAEKGCIMNFSTGPPFMVFSKSRKDEASLEKMGMCLEFNL 413
Qy 467 PQNPIQLPGLDLPGTSYDANROCOFTFGEDSKHCPDAASTCTLWCTGTGTGGVLVCQTKHF 526
Db 414 PEVQAQFG-----RKCGNGYVEEGEC----- 436
Qy 527 PWADGTSGBEGKWCINGKCVKN-----HRKHFDTPFHGSMGWMGCPWGDSCRTC 575
Db 437 -----DCGEPEECTN-RCCNATTTCLKPDAVCAH-----GQCCEDC 471
Qy 576 -----GGGVQVYTMRECDNPVPKNGKYCEGRKRVYRSCNLEBDCPDN----- 616
Db 472 QLKPPGTACRGSSNSCDLP-----EFCTG-----TAPHC PANVYLHDGHPCOQVDG 517
Qy 617 ---NGKTFREEQBNAHNEFKASFGSGPAVEWIPKYAGVSPKCKLIKQAKGIYGFVL 673
Db 518 YCYNG-----ICQTHEQOCQVTLWGPAGAK-----PAPGICFERVN 551
Qy 674 QPKVVDGTPCSPDSTSVQVQOCVKAGCDRIIDSKKFKDFKCGVCGNGSTCKKISGSVTS 733
Db 552 SA-----GDP-----YGNCK-----DSKSAFAKCEL---RDAKCGKIQCOQGA 587
Qy 734 AKPGYHDIITPTGATNIEVKORNGSRNNGSFLAIAAADCTYILNGDYTLSTLEQDIM 793
Db 588 SRP-----VIGTNVAVSIETNIPQEGGR-----ILCRGTHVYLG-----DMP 625
Qy 794 YKGVVLRYSGSSAALERI 811
Db 626 DPGILVL--AGTKCAEGKI 641

RESULT 14
S48169
metalloproteinase (EC 3.4.24.-) H-II precursor - carpet viper (fragment)
N:Contains: disintegrin
C:Species: Echis pyramidum leakeyi
C:Date: 06-Dec-1996 #sequence_revision 06-Dec-1996 #text_change 09-Jul-2004
R:Paine, M.J.I.; Moura-Da-Silva, A.M.; Theakston, R.D.G.; Crampton, J.M.
Eur. J. Biochem. 224, 483-488, 1994
A:Title: Cloning of metalloproteinase genes in the carpet viper (Echis pyramidum leakeyi).
A:Reference number: S48169; MUID:95010025; PMID:7925363
A:Accession: S48169
A:Molecule type: mRNA
A:Residues: 1-549 <PAI>
C:Cross-references: UNIPROT:Q90500; EMBL:X78971; NID:G763094; PIDN:CAA55566.1; PID:G76309
C:Superfamily: mouse meltin alpha; disintegrin homology
C:Keywords: hydrolase; metalloproteinase; venom; zinc
F:1-138/Domain: propeptide (fragment) #status predicted <PRO>
F:139-549/Product: metalloproteinase H-II #status predicted <MAT>
F:350-432/Domain: disintegrin homology <Dis>
F:284/Active site: Glu #status predicted

Query Match 6.6%; Score 347; DB 2; Length 549;
Best Local Similarity 23.0%; Pred. No. 1.5e-15;
Matches 152; Conservative 80; Mismatches 206; Indels 222; Gaps 32;

Qy 106 GRKSGSDTLPETDLAHCFYSGTVNGDPSSAAALSLCEGVRGAFYLLGEAYFIQPLPAAS 165
```

Db 38 GREITNPVED-----HCYHGRVQNDASHSSASISACNGLKGFKLQGETYFIEPLKI--- 91
Qy 166 ERLATAAPCEKPPAPLQFHLRRNRQDVGCGTVVDDPRPTGKAEDEDEDEGEDE 225
Db 92 -----PDSEAHAYKYENTEKEDQ 110
Qy 226 GPO-----WSPQDPALQGVQPTGTSIRKRFVSSHRYVETMLVADQSMAPFHG--- 275
Db 111 APQMGCVTHTNMESDEPIKEA---SRLVASSEQQSYDNFRYVKLVIVVDRHMVTYKND 167
Qy 276 SGLKHYLLTLFSAARLYKHPSTRNSVSLVVKI-----LVTHDEQKGEVTSNAALTIR 330
Db 168 STIRTRIYEMVNTVNEIYTHAIR--VALVLEFWSNGDLI-----NVTSAEHTLN 217
Qy 331 NFCNMOKQHPSPD-RDAEHYDTAILFTQDLGSGTCTDLGMADVGTVCDSRSCSVIE 389
Db 218 LFGVWR-----ASDLLSRHRHDNAHLLTAIDLNG-----FTIGLRDVSSWCQATRSVGVVQ 268
Qy 390 DDG---LQAAFTTAHELGHVFNMPHDDAKQOCASLNGVNDQSHMMASMLSNLDHSPQWSPC 446
Db 269 DHSPTVRVAVTMAHEMGNLGMH--DGNHC-----NCGANSICMAAVLRN-PAPEYFSDC 322
Qy 447 SGYMTSFLDNGHGBCLMDKPNQPIQLPGDLPGTSVDANRQCFTEGDSKSCPDAASTC 506
Db 323 SRRYQNFLTNYPDCITIRPK-----TDIVSPQVCNGGLLDEBEC--DCGSPA 371
Qy 507 STLWCTGTSGGVLVCQTKHPFADGTSCEGKW--CINGKCVNKNHRKHFDPFHGSGM 564
Db 372 N-----CO---YPCDAASCKLHNSVECFHCCDQCFKPACTECRG----- 411
Qy 565 WGPWGDSCSRTCGGVQVYTMRECDNPVPKNGKYCEG-----KRVRYRSCNLE 611
Db 412 -----IRSECDLP-----EYCTGQSAECLRMFSTRKWTMPKLSLLIQ 448
Qy 612 -DCPDNNGKTFREOCEAHNEFSKASFGSGPAVEWIPKYGSPKDRC-KLICQAKGIGY 669
Db 449 WTCP-----IMGYQCYAH-----FGQNAVY-----GQDACFEINKEGKGFY 485
Qy 670 -----PFVLQPK-----VVD-GTPCSPDSTSVCVQGCQV 697
Db 486 CRKENDVPIPCAQEDVKCGRLCFETEPNMCPRYPGDEGWDPTGTC--EDKKVCINGKCI 543

RESULT 15
S24789

jararagin C precursor - jararaca (fragment)
N:Alternate names: single chain botrocetin
N:Contains: disintegrin-like 28k protein; hemorrhagic proteinase (EC 3.4.24.-)
C:Species: Bothrops jararaca (jararaca)
C>Date: 20-Feb-1995 #sequence revision 29-Aug-1997 #text_change 09-Jul-2004
C:Accession: S24789; J2245; A4463; A37958; J22373
R:Paine, M.J.I.
submitted to the EMBL Data Library, August 1992
A:Reference number: S24789
A:Accession: S24789
A:Molecule type: mRNA
A:Residues: 1-571 <PAI>
A:Cross-references: UNIPROT:P30431; EMBL:X68251; NID:g62467; PID:g62468
R:Usami, Y.; Fujimura, Y.; Miura, S.; Shima, H.; Yoshida, E.; Yoshioka, A.; Hirano, K.;
Biochem. Biophys. Res. Commun. 201, 331-339, 1994
A:Title: A 28 kDa-protein with disintegrin-like structure (jararagin-C) purified from B
A:Reference number: J2245; MUID:94256999; PMID:8198592
A:Accession: J2245
A:Molecule type: protein
A:Residues: 360-571 <USA>
A:Experimental source: venom
R:Paine, M.J.; Desmond, H.P.; Theakston, R.D.; Crampton, J.M.
J. Biol. Chem. 267, 22869-22876, 1992
A:Title: Purification, cloning, and molecular characterization of a high molecular weight
ily.
A:Reference number: A4463; MUID:93054601; PMID:1385408
A:Accession: A4463

A:Molecule type: mRNA
A:Residues: 1-23,'Q','25-92,'G','94-131,'G','133-169,'Q','171-571 <PA2>
A:Cross-references: GB:X68251
A:Experimental source: venom gland
A:Note: sequence inconsistent with the nucleotide translation
R:Fujimura, Y.; Tican, K.; Usami, M.; Suzuki, M.; Oyama, R.; Matsui, T.; Fukui, H.; Sug
Biochemistry 30, 1957-1964, 1991
A:Title: Isolation and chemical characterization of two structurally and functionally di
A:Reference number: A37958; MUID:91129280; PMID:1993206
A:Accession: A37958
A:Molecule type: protein
A:Residues: 360-372,'E','374-378,'X','380-384 <FUI>
A:Note: 361-Val was also found
C:Comment: Inhibits collagen- and ADP-induced platelet aggregation.
C:Superfamily: mouse metrin alpha; disintegrin homology
C:Keywords: hydrolase; metalloproteinase; venom; zinc
F:360-571/Product: jararagin C #status experimental <MA>
F:362-444/Domain: disintegrin homology <DIS>
F:295,299,305/Binding site: zinc (His) #status predicted
F:296/Active site: Glu #status predicted

Query Match 6.5%; Score 343; DB 2; Length 571;

Best Local Similarity 21.9%; Pred. No. 3e-15;

Matches 158; Conservative 94; Mismatches 228; Indels 242; Gaps 36;

Qy 48 DALGRPSDEELVPELERVPGHGTTRLRLHAFDQDLDDVPPDSSFLAPGFTLQNVGR 107

Db 14 DAMQVEFKVNGEPPVHLKNGK-----LFSKDYs-EIHVSPD-----GR 52

Qy 108 KSGSTPTLPETDLAHCFSYGTVNGDPSSAAALSCEGVGAFYLLGEAYFTQPLPAASER 167

Db 53 EITTYPPVED-----HCYHGRINDADSTASACNGLKGFYKQRETYFIEPLK----- 104

Qy 168 LATAAPGEKPPAPLQFHLRRNRQDVGCGTVVDDPRPTGKAEDEDEGEDEGEP 227

Db 105 -----PDSEAHAYKYENVEKEDAP 125

Qy 228 Q-----WSPQDPALQGVQPTGTSIRKRFVSSHRYVETMLVADQSMAPFHGSLKH 280

Db 126 KMGCVTQNKWSEY-IKKASQLAFTA--EQRY-DPYKIBFFVVVDQGTVTKNNGDLD- 180

Qy 281 YLLTFLSVAARLYKHPSTRNSVSLVVKLVTHDEQKGEVTSN-----AALTNRN 331

Db 181 -----KIKARMYELANIVNE-----IFRYLMHVVALVLEIWSNGDKITVKPDDVDTLNS 230

Qy 332 FCNMOKQHPSPD-RDAEHYDTAILFTQDLGSGTCTDLGMADVGTVCDSRSCSVIED 390

Db 231 FAEWRK-----TDLLTRKHNDAQLLTAIDFNG-----PTIGYAYIGSMCHPKRSYGVQD 281

Qy 391 ---DGLQAAFTTAHELGHVFNMPHDDAK-QCASLNGVNDQSHMMASMLSNLDHSPQWSPC 446

Db 282 YSPINLVAVIMAHMGNLGIHHDGTGSCSGDYPCI-----MGPTISN-EPSKFFSNC 334

Qy 447 SGYMTSFLDNGHGBCLMDKPNQPIQLPGDLPGTS-YDANRQCFTEGDSKH-CPDAAS 504

Db 335 SYIQCWDFIMNHNPECIINEPLGTDIISPPVCGNLELVEBECDCGTPEQCNECCDAAT 394

Qy 505 TCSTLWCTGTSGGVLVCQTKHPFADGTSCEBGKWCINGKCVNKNHRKHFDPFHGSGM 564

Db 395 -----CKLK-----SGSQGH----- 405

Qy 565 WGPWGDSCSRTC-----GGGVQVYTMRECDNPVPKNGKYCEGKRVYRSC-----NLEED 612

Db 406 ---GDCCEQCKFSKSGTECRASMSECD-----FAEHCTGS-----SECPADVPHKNGQP 452

Qy 613 CPDNNKGTFR-----EEQCEAHNEFSKASFGSGPAVEWIPKYG 651

Db 453 CLDNYGYCNGNCPIMYHCYALFGADVYEADSCFDKNQ-----KGNYYGYCKRENG 505

Qy 652 ----VSPKD-RC-KLICQAKGIG-----YFVLQPK-----VVDGTPCSPDSTSVCVQGO 695

Db 506 KKIPCAPEDVCGRLYCKDNSPGQNPCKMFFSYNDDEHKGMVLPGTKCA--DGKVCNSGH 563

Qy 696 CV 697
Db 564 CV 565

Search completed: August 22, 2005, 10:19:38
Job time : 53 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2005, 10:06:04 ; Search time 179 Seconds
(without alignments)
2766.372 Million cell updates/sec

Title: US-09-989-687-126
Perfect score: 5287
Sequence: 1 MORAVPEGFRKLGSDMGN.....CDPLKPKHFIDFCTMAECS 967

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_03.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	5232	99.0	967	1	AT51_HUMAN	Q9ubi8 homo sapien
2	5234	98.8	967	2	Q8NE26	Q8ne26 homo sapien
3	4293.5	81.2	968	1	AT51_MOUSE	P97857 mus musculus
4	4293	81.2	967	2	Q68EJ2	Q68ej2 rattus norv
5	4287	81.1	967	1	AT51_RAT	Q9wug1 rattus norv
6	3902.5	73.8	759	2	Q8RZM8	Q8hzm8 equus cabal
7	2477.5	46.9	950	1	AT15_HUMAN	Q8te58 homo sapien
8	2320.5	43.9	890	1	AT58_HUMAN	Q9up79 homo sapien
9	2265.5	42.9	905	1	AT58_MOUSE	P57110 mus musculus
10	2124	40.2	837	1	AT54_HUMAN	Q75173 homo sapien
11	2117	40.0	837	2	Q6UW8	Q6uwa8 homo sapien
12	2106	39.8	837	2	Q6P4Q8	Q6p4q8 homo sapien
13	2105	39.8	833	2	Q8K384	Q8kj384 mus musculus
14	2104.5	39.8	845	2	Q8BNJ2	Q8bnj2 mus musculus
15	2075.5	39.3	839	2	Q7YS95	Q7ys95 bos taurus
16	1979	37.4	630	1	AT54_RAT	Q9esp7 rattus norv
17	1976	37.4	893	2	Q6A017	Q6a017 mus musculus
18	1924	36.4	930	1	AT55_HUMAN	Q9una0 homo sapien
19	1911.5	36.2	930	1	AT55_MOUSE	Q9r001 mus musculus
20	1904.5	36.0	928	2	Q6TY19	Q6ty19 rattus norv
21	1862	35.2	1306	1	AT20_MOUSE	P59511 mus musculus
22	1825	34.5	1935	1	AT59_HUMAN	Q9p2n4 homo sapien
23	1794.5	33.9	867	2	Q6SKM3	Q66km3 xenopus tro
24	1727.5	32.7	1911	1	AT20_HUMAN	P59510 homo sapien
25	1672	31.6	623	2	Q8BGP4	Q8bgp4 m mus muscu
26	1614	30.5	562	1	AT15_MOUSE	P59384 mus musculus
27	1598	30.2	2165	2	Q19751	Q19791 caenorhabdi
28	1359	25.7	1641	2	Q68SA9	Q68sa9 mus musculus
29	1321	25.0	1886	2	Q6P7J9	Q6p7j9 homo sapien
30	1280	24.2	1593	1	AT12_HUMAN	P58397 homo sapien
31	1279.5	24.2	1221	2	Q6P4R5	Q6p4r5 homo sapien

32	1277.5	24.2	1224	1	AT16_HUMAN	Q8te57 homo sapien
33	1277	24.2	1233	2	Q69Z28	Q69z28 mus musculus
34	1276.5	24.1	1070	2	Q8CG28	Q8cg28 mus musculus
35	1273	24.1	1077	1	AT10_HUMAN	Q9h324 homo sapien
36	1273	24.1	1081	1	AT18_HUMAN	Q8te60 homo sapien
37	1267.5	24.0	1009	2	Q8BKV1	Q8bky1 m mus muscu
38	1267.5	24.0	1600	2	Q811B3	Q81lb3 mus musculus
39	1252.5	23.7	1059	2	Q9W493	Q9w493 drosophila
40	1248	23.6	1537	2	Q7KSH7	Q7ksh7 drosophila
41	1248	23.6	1688	2	Q8SXB0	Q8sxb0 drosophila
42	1217.5	23.0	997	1	AT57_HUMAN	Q9ukp4 homo sapien
43	1184.5	22.4	988	2	Q7PWY7	Q7pwy7 anopheles g
44	1184	22.4	1092	2	Q8BKA1	Q8bka1 mus musculus
45	1158	21.9	900	2	Q8K206	Q8k206 mus musculus

ALIGNMENTS

RESULT 1
AT51_HUMAN
ID AT51_HUMAN STANDARD; PRT; 967 AA.
AC Q9UHf8; Q9NSJ8; Q9P2K0; Q9UH83; Q9UP80;
DT 30-MAY-2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE ADAMTS-1 precursor [EC 3.4.24.-] (A disintegrin and metalloproteinase
DE with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1) (METH-1).
GN Name=ADAMTS1; Synonyms=KIAA1346, METH1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Casas C., Pritchard M.A., Estivill X., Arbones M.L.;
RT "Cloning, characterization and mapping on human chromosome 21 of the
RT orthologue of murine Adamts-1.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A., AND FUNCTION.
RC TISSUE=Heart;
RX MEDLINE=99367466; PubMed=10438512; DOI=10.1074/jbc.274.33.23349;
RA Vazquez F., Hastings G., Ortega M.-A., Lane T.F., Oikemus S.,
RA Lombardo M., Iruela-Arispe M.L.;
RT "METH-1, a human ortholog of ADAMTS-1, and METH-2 are members of a new
RT family of proteins with angio-inhibitory activity.";
RL J. Biol. Chem. 274:23349-23357(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Endothelial cells;
RX MEDLINE=20247184; PubMed=10785405;
RA Glienke J., Schmitt A.O., Pilarsky C., Hinzmann B., Weiss B.,
RA Rosenthal A., Thierach K.H.;
RT "Differential gene expression by endothelial cells in distinct
RT angiogenic states.";
RL Eur. J. Biochem. 267:2820-2830(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20181126; PubMed=10718198;
RA Nagase T., Kikuno R., Ishikawa K.-I., Hirosewa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XVI.
RT The complete sequences of 150 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 7:65-73(2000).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=20289799; PubMed=10830953; DOI=10.1038/35012518;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,
RA Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,

RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
 RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
 RA Shintani A., Sasaki T., Nagamine K., Mitsuyma S., Antonarakis S.E.,
 RA Minoshima S., Shimizu N., Nordlie G., Hornischer K., Brandt P.,
 RA Scharf M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H.,
 RA Ramer J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,
 RA Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
 RA Lehrach H., Reinhardt R., Vaspou M.-L.; 21."
 RT "The DNA sequence of human chromosome 21."
 RL Nature 405:311-319(2000).
 RN [6]
 RP SEQUENCE OF 418-967 FROM N.A.
 RC TISSUE-Melanoma;
 RA Blum H., Bauersachs S., Mewes H.-W., Weil B., Wienann S.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be
 CC involved in its turnover (By similarity). Has angiogenic inhibitor
 CC activity. Active metalloprotease, which may be associated with
 CC various inflammatory processes as well as development of cancer
 CC cachexia. May play a critical role in follicular rupture.
 CC -!- CATALYTIC ACTIVITY: Cleaves aggrecan at the 1938-Glu-|-Leu-1939
 CC site, within the chondroitin sulfate attachment domain.
 CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
 CC matrix (By similarity).
 CC -!- DOMAIN: The spacer domain and the TSP type-1 domains are important
 CC for a tight interaction with the extracellular matrix.
 CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By
 CC similarity).
 CC -!- SIMILARITY: Belongs to the peptidase M12B family.
 CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
 CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF170084; AAF15317.1; -;
 CC EMBL; AF060152; AAD48080.1; ALT_INIT.
 CC EMBL; AF207664; AAF23772.1; -;
 CC EMBL; AB037767; BAA32584.1; ALT_INIT.
 CC EMBL; AP001697; BAA95502.1; -;
 CC EMBL; AL162080; CAB82413.1; -;
 CC PIR; T47158; T47158.
 CC HSP; P07996; 1LSL.
 CC MEROPS; M12.22; -;
 CC Genew; HGNC:217; ADAMTS1.
 CC H-InvDB; HIX0016042; -;
 CC MIM; 605174; -;
 CC GO; GO:0008237; F:metallopeptidase activity; TAS.
 CC GO; GO:0007229; P:integrin-mediated signaling pathway; TAS.
 CC GO; GO:0008285; P:negative regulation of cell proliferation; TAS.
 CC InterPro; IPR006586; ADAM_cysteine.
 CC InterPro; IPR010294; ADAM_spacer1.
 CC InterPro; IPR001762; Disintegrin.
 CC InterPro; IPR001590; Peptidase M12B.
 CC InterPro; IPR002870; Peptidase M12B.
 CC InterPro; IPR006025; Pept_M_Zn_BS.
 CC InterPro; IPR000884; TSP1.
 CC InterPro; IPR000885; TSP1.
 CC Pfam; PF05986; ADAM_spacer1; 1.
 CC Pfam; PF01562; Pep_M12B_propep; 1.
 CC Pfam; PF01421; Reprolysin; 1.
 CC Pfam; PF00090; TSP1; 3.
 CC PRINTS; PR01705; TSP1REPEAT.
 CC SMART; SM00608; ACR; 1.
 CC SMART; SM00209; TSP1; 3.
 CC PROSITE; PS0215; ADAM_MEPRO; 1.
 CC PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.

DR PROSITE; PS0214; DISINTEGRIN_2; FALSE_NEG.
 DR PROSITE; PS0092; TSP1; 3.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW Extracellular matrix; Glycoprotein; Heparin-binding; Hydrolase;
 KW Metalloprotease; Repeat; Signal; Zinc; Zymogen.
 FT SIGNAL 1 49 Potential.
 FT PROPEP 50 252 By similarity.
 FT CHAIN 253 967 ADAMTS-1.
 FT DOMAIN 253 475 Metalloprotease.
 FT DOMAIN 476 559 Disintegrin-like.
 FT DOMAIN 559 614 TSP type-1 1.
 FT DOMAIN 617 724 Cys-rich.
 FT DOMAIN 725 849 Spacer.
 FT DOMAIN 854 905 TSP type-1 2.
 FT DOMAIN 908 967 TSP type-1 3.
 FT DOMAIN 843 846 Poly-Lys.
 FT SITE 198 198 Cysteine switch (Potential).
 FT METAL 401 401 Zinc (catalytic) (By similarity).
 FT ACT SITE 402 402 By similarity.
 FT METAL 405 405 Zinc (catalytic) (By similarity).
 FT METAL 411 411 Zinc (catalytic) (By similarity).
 FT CARBOHYD 547 547 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 720 720 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 764 764 N-linked (GlcNAc...) (Potential).
 FT CONFLICT 227 227 P -> A (in Ref. 4 and 5).
 FT CONFLICT 468 468 Q -> H (in Ref. 1).
 FT CONFLICT 561 561 S -> N (in Ref. 1).
 SQ SEQUENCE 967 AA; 105383 MW; C18938932474LED1 CRC64;
 Query Match 99.0%; Score 5232; DB 1; Length 967;
 Best Local Similarity 99.0%; Pred. No. 0;
 Matches 957; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
 Qy 1 MQRVAVPGFRRKLGSDMGNAERAPGSRSGPVPVTLTLLAALLAVSDALGRPSEDEEL 60
 Db 1 MQRVAVPGFRRKLGSDMGNAERAPGSRSGPVPVTLTLLAALLAVSDALGRPSEDEEL 60
 Qy 61 VPELVRVGHGTRLRHLHAFDQDLDPDSSFLAPGFTLQNVGRKSGDPLPETDL 120
 Db 61 VPELVRVGHGTRLRHLHAFDQDLDPDSSFLAPGFTLQNVGRKSGDPLPETDL 120
 Qy 121 AHCFYSGTVNGDPSSAAALSICGVGAFYLLGAYFIQPLPAASERLATAAPGKPPAP 180
 Db 121 AHCFYSGTVNGDPSSAAALSICGVGAFYLLGAYFIQPLPAASERLATAAPGKPPAP 180
 Qy 181 LQFHLLRRNQDVGTCGVVDDEPRPTGKAETDEDEGEDEGEDEGEDEGEDEGEDEGE 240
 Db 181 LQFHLLRRNQDVGTCGVVDDEPRPTGKAETDEDEGEDEGEDEGEDEGEDEGEDEGE 240
 Qy 241 QPTGTGSIKRRFVSSHRYVETMLVADQSMAPFHGSLKHVLLTLFVVAARLYKHPISRN 300
 Db 241 QPTGTGSIKRRFVSSHRYVETMLVADQSMAPFHGSLKHVLLTLFVVAARLYKHPISRN 300
 Qy 301 SVSLVWVKILVIHDEQKGPVTSNAALTNRNFCNQKQHNPPSDRDAEHYDTAILFTQD 360
 Db 301 SVSLVWVKILVIHDEQKGPVTSNAALTNRNFCNQKQHNPPSDRDAEHYDTAILFTQD 360
 Qy 361 LCQSQTCDTLGMADVGTVCPSRSCSVIEDDGLQAAFTTAHELGHVFMNPHDDAKQCASL 420
 Db 361 LCQSQTCDTLGMADVGTVCPSRSCSVIEDDGLQAAFTTAHELGHVFMNPHDDAKQCASL 420
 Qy 421 NGVNQDSHMMASMLSNLDHSPQSPSCSGYMTSFLDNGHGECMLDKPQNPTQLPGDLPGT 480
 Db 421 NGVNQDSHMMASMLSNLDHSPQSPSCSGYMTSFLDNGHGECMLDKPQNPTQLPGDLPGT 480
 Qy 481 SYDANRQCQFTFGEDSKHCPDAAASTCTLWCTGTSGGLVLCQTKHFPWADGTCGEGKWC 540
 Db 481 SYDANRQCQFTFGEDSKHCPDAAASTCTLWCTGTSGGLVLCQTKHFPWADGTCGEGKWC 540
 Qy 541 INKCVNKNHRKHFDPTPFHGSWGHWGPDGCSRTCGGQVQYTWRECDNPPVKNQKGYCEG 600
 Db 541 INKCVNKNHRKHFDPTPFHGSWGHWGPDGCSRTCGGQVQYTWRECDNPPVKNQKGYCEG 600

QY 661 IQOAKGIGYFFVLPQKVDGTPCSPDSTSVCGQCVKAGCDRIIDSKKKFKGCVCGN 720
 |||||
 Db 661 IQOAKGIGYFFVLPQKVDGTPCSDSTSVCGQCVKAGCDRIIDSKKKFKGCVCGN 720
 |||||
 QY 721 GSTCKKISGVSPTSAPKGYHDIITIPGATNIEVKQNRGSRNNGSFLAIKAADGTIYN 780
 |||||
 Db 721 GSTCKKISGVSPTSAPKGYHDIITIPGATNIEVKQNRGSRNNGSFLAIKAADGTIYN 780
 |||||
 QY 781 GDTLTSTLEODIMYGVNLVSGSSAALERIRSFSPLEKPLTIQVLTGVALRPPKIKYTY 840
 |||||
 Db 781 GDTLTSTLEODIMYGVNLVSGSSAALERIRSFSPLEKPLTIQVLTGVALRPPKIKYTY 840
 |||||
 QY 841 FVKKKESFNALPTFSANWIEBEGSCSKELGWORRLVECRDINGQPAASECAKEVKPAS 900
 |||||
 Db 841 FVKKKESFNALPTFSANWIEBEGSCSKELGWORRLVECRDINGQPAASECAKEVKPAS 900
 |||||
 QY 901 TRPCADHPCPQWLGEWSCKTCGKYKTKSLKCLSHDGGVLSHSDCDPLKKPKHFIDF 960
 |||||
 Db 901 TRPCADHPCPQWLGEWSCKTCGKYKTKSLKCLSHDGGVLSHSDCDPLKKPKHFIDF 960
 |||||
 QY 961 CTMAECS 967
 |||||
 Db 961 CTMAECS 967
 |||||

RESULT 3
 ATSL_MOUSE
 ID ATSL_MOUSE STANDARD; PRT; 968 AA.
 AC P97857; O54768;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE ADAMTS-1 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
 RE with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1).
 GN Name=Adamts1;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SVJ;
 RX MEDLINE=98110583; PubMed=9441751; DOI=10.1006/geno.1997.5064;
 RA Kuno K., Lizasa H., Ohno S., Matsushima K.;
 RT "The exon/intron organization and chromosomal mapping of the mouse
 ADAMTS-1 gene encoding an ADAM family protein with TSP motifs.";
 RL Genomics 46:466-471(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97150761; PubMed=8995297; DOI=10.1074/jbc.272.1.556;
 RA Kuno K., Kanada N., Nakashima E., Fujiki F., Ichimura F.,
 RA Matsushima K.;
 RT "Molecular cloning of a gene encoding a new type of metalloproteinase-
 disintegrin family protein with thrombospondin motifs as an
 RT inflammation associated gene.";
 RL J. Biol. Chem. 272:556-562(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22380257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.D., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.L., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan F.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Heiton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP CHARACTERIZATION, AND MUTAGENESIS OF GLU-403.
 RX MEDLINE=99303657; PubMed=10373500; DOI=10.1074/jbc.274.26.18821;
 RA Kuno K., Terashima Y., Matsushima K.;
 RT "ADAMTS-1 is an active metalloproteinase associated with the
 RT extracellular matrix.";
 RL J. Biol. Chem. 274:18821-18826(1999).
 RN [5]
 RP FUNCTION.
 RX MEDLINE=20389568; PubMed=10930576; DOI=10.1016/S0014-5793(00)01854-8;
 RA Kuno K., Okada Y., Kawashima H., Nakamura H., Miyasaka M., Ohno H.,
 RA Matsushima K.;
 RT "ADAMTS-1 cleaves a cartilage proteoglycan, aggrecan.";
 RL FEBS Lett. 478:241-245(2000).
 RN [6]
 RP FUNCTION, AND INDUCTION.
 RX MEDLINE=20243757; PubMed=10781075; DOI=10.1073/pnas.080073497;
 RA Robker R.L., Russell D.L., Espey L.L., Lydon J.P., O'Malley B.W.,
 RA Richards J.S.;
 RT "Progesterone-regulated genes in the ovulation process: ADAMTS-1 and
 RT cathepsin L proteases.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:4689-4694(2000).
 CC -!- FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be
 CC involved in its turnover. Has angiogenic inhibitor activity (By
 CC similarity). Active metalloproteinase, which may be associated with
 CC various inflammatory processes as well as development of cancer
 CC cachexia. May play a critical role in follicular rupture (By
 CC similarity).
 CC -!- CATALYTIC ACTIVITY: Cleaves aggrecan at the 1691-Glu-|-Leu-1692
 CC site, within the chondroitin sulfate attachment domain.
 CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
 CC matrix.
 CC -!- INDUCTION: Induced in vitro in colon adenocarcinoma cells by
 CC interleukin-1, or in vivo in kidney and heart by
 CC lipopolysaccharide. Also induced by LH stimulation in granulosa
 CC cells of preovulatory follicles.
 CC -!- DOMAIN: The spacer domain and the TSP type-1 domains are important
 CC for a tight interaction with the extracellular matrix.
 CC -!- PTM: The precursor is cleaved by a furin endopeptidase.
 CC -!- SIMILARITY: Belongs to the peptidase M12B family.
 CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
 CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
 CC -!- CAUTION: Ref.2 sequence differs from that shown due to a
 CC frameshift in position 7.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB001735; BAA24501.1; ALT INIT.
 DR EMBL; D67076; BAA11088.1; ALT_FRAME.
 DR EMBL; BC040382; AAH40382.1; -.
 DR EMBL; BC050834; AAH50834.1; -.
 DR HSPB; P07996; ILSL.
 DR MEROPS; M12.222; -.
 DR MGD; MGI:109249; Adamts1.
 DR InterPro; IPR010294; ADAM_spacer1.
 DR InterPro; IPR001762; Disintegrin.
 DR InterPro; IPR006025; Pept M. Zn BS.
 DR InterPro; IPR001590; Peptidase_M12B.

```

DR InterPro; IPR002870; Peptidase_M12B_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR000885; TSP 1.
DR Pfam; PF05986; ADAM spacer1; 1.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR Pfam; PF01421; Repolysin; 1.
DR Pfam; PF00090; TSP1; 3.
DR PRINTS; PR01705; TSP1REPEAT.
DR PROSITE; PS00215; ADAM_MEPRO; 1.
DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR PROSITE; PS0214; DISINTEGRIN_2; FALSE_NEG.
DR PROSITE; PS0092; TSP1; 3.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR Extracellular matrix; Glycoprotein; Heparin-binding; Hydrolase;
KW Metalloprotease; Repeat; Signal; Zinc; Zymogen.
FT SIGNAL 1 48
FT PROPEP 49 253
FT CHAIN 254 968
FT DOMAIN 254 476
FT DOMAIN 477 559
FT DOMAIN 560 615
FT DOMAIN 618 725
FT DOMAIN 726 850
FT DOMAIN 855 911
FT DOMAIN 912 968
FT SITE 195 199
FT SITE 206 206
FT METAL 402 402
FT ACT_SITE 403 403
FT METAL 406 406
FT METAL 412 412
FT CARBOHYD 548 548
FT CARBOHYD 721 721
FT CARBOHYD 765 765
FT CARBOHYD 783 783
FT CARBOHYD 946 946
FT MUTAGEN 403 403
FT CONFLICT 335 335
FT CONFLICT 425 425
SQ SEQUENCE 968 AA; 105841 MW; 42EBDA55499FB6C1 CRC64;

```

Query Match 81.2%; Score 4293.5; DB 1; Length 968;

Best Local Similarity 81.0%; Pred. No. 5.5e-270;

Matches 790; Conservative 58; Mismatches 112; Indels 15; Gaps 5;

```

Qy 1 MORAVPEGFRGRKLGSDMGNAERAPGRSFGVPVTLTLLAAA---LLAVSDALGRPSEED 57
Db 1 MOPKVPGLGRKQKPCSDMGDVORARSRGSLSAHMLLLLLLSITMLLCARGHGRPTBED 60
Qy 58 BELVVPLELERVGHG--TTRLRLHAFDQDLDDVPPDSSFLAPGFTLQNVGRKSGSDTPL 115
Db 61 BELVLPSLERAPGHDSTTTRLRLDAFGQQLHLKLPDSDGFLAPGFTLQTVGRSPGSEAQH 120
Qy 116 --PETDLAHCFVSGTVNGDPSSAALSLCEGVRGAFYLLGEAYFTQPLP-AASEBLATAA 172
Db 121 LDPTGDLAHCFVSGTVNGDPGSAALSLCEGVRGAFYLLQGEFFITQAPAGVATERLAPAV 180
Qy 173 PGEKPPAPLQPHLLRRNRQDVGCTGVGVDDPRPTGKAETDEDEGTEGEDEGPQWSPQ 232
Db 181 PEESSARPPPHILRRRRRGSGAGKGVNDDTLPT-----SDSRPESQNRQWNPVR 233
Qy 233 DPALQGVGOPTGTGIRKKRFVSSHRYVETMLVADQSMAPHEGSLGKHYLLTLFVSAARL 292
Db 234 DTPPDACKSPGSGIRKKRFVSSPRYVETMLVADQSMADFHGSLGKHYLLTLFVSAARF 293
Qy 293 YKHPISIRNSVLVVKILVHDEQGPVTSNAALTFLNFCNQKHNPSPDRDAEHYDT 352
Db 294 YKHPISIRNSISLVVKILVIEEQGPVTSNAALTFLNFCNQKHNPSPDRDPEHYDT 353
Qy 353 ALLFTRODLCSQTCDTLGLMADVGVTCOPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHD 412
Db 354 ALLFTRODLCSQTCDTLGLMADVGVTCOPSRSCSVIEDDGLQAAFTTAHELGHVFNMPHD 413

```

RESULT 4

```

Q68EJ2 ID Q68EJ2 PRELIMINARY; PRT; 967 AA.
AC Q68EJ2;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE A disintegrin and metalloproteinase with thrombospondin motifs 1.
GN Name=Adams1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McWeary K.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

```

```

Qy 413 DAKOCASLNGVNDQSHMMASMLSNLDHSQWPCSGYMTSFLDNGHGECGLMDKPNPIQ 472
Db 414 DAKHCASLNGVTGDSLHMASMLSLDHSQWPCSGYMTSFLDNGHGECGLMDKPNPIK 473
Qy 473 LPGLDPTGSYDANRQCQPTFCGEDSKHCPDAASTCTLWCTGTSGGLVLCQTKHPFWADGT 532
Db 474 LPSDLPGTLYDANRQCQPTFCGESKHCPDAASTCTTLLWCTGTSGGLVLCQTKHPFWADGT 533
Qy 533 SCGEGKWCINGKCVNKNRHRKFDTPFHGSGWGMGPDGSCRTCGGQVQYTMRECDNPVPK 592
Db 534 SCGEGKWCVSGKCVNKNRHRKFDTPFHGSGWGMGPDGSCRTCGGQVQYTMRECDNPVPK 593
Qy 593 NGKYCEGKRVYRSCNLEDPCDNNNGKTFREQCEAHNEFASKAFSGSPAWEVLPKYAGV 652
Db 594 NGKYCEGKRVYRSCNIEDPCDNNNGKTFREQCEAHNEFASKAFSGSPAWEVLPKYAGV 653
Qy 653 SPKDRCKLICAKGIGYFFVLQPKVVDGTPCSPDSTSVCVQGCVKAGCDRIIDSKKKFD 712
Db 654 SPKDRCKLTCEAKGIGYFFVLQPKVVDGTPCSPDSTSVCVQGCVKAGCDRIIDSKKKFD 713
Qy 713 KCGVCGNGSTCKKISGSVTSAPGYHDIITPTGATNIEVKQRNQRNGSRNGSFLAIKA 772
Db 714 KCGVCGNGSTCKKMGIVTSTRPGYHDIITPTGATNIEVKQRNQRNGSRNGSFLAIKA 773
Qy 773 ADGYIILNGDVTLTSLDODIMYKGVLYRYSOSSAALERISFSPKLEPLTTIQLVTVGNAL 832
Db 774 ADGYIILNGDVTLTSLDODIMYKGVLYRYSOSSAALERISFSPKLEPLTTIQLVTVGNAL 833
Qy 833 RPKIKYTYFVKKKESFNAIFTSANWIEEWGECSCELGWQRRLVBCRDINGOPASEC 892
Db 834 RPKIKYTYFVKKKESFNAIFTSANWIEEWGECSCELGWQRRLVBCRDINGOPASEC 893
Qy 893 AKVKPASTRCPADHPCPQWOLGSEWSSCKTCGKGYKTKSLKSHDGGVLSHSDCDPLK 952
Db 894 AKVKPASTRCPADHPCPQWOLGSEWSSCKTCGKGYKTKSLKSHDGGVLSHSDCDPLK 953
Qy 953 KPKHIFDFTWAECS 967
Db 954 KPKHYIDFTLTQCS 968

```

RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Director MSC Project;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC080233; AAH0237.1; -;
DR GO; GO:0007229; P: integrin-mediated signaling pathway; IEA.
DR InterPro; IPR006586; ADAM cysteine.
DR InterPro; IPR010294; ADAM spacer1.
DR InterPro; IPR001590; Peptidase M12B.
DR InterPro; IPR002870; Peptidase M12B_N.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP 1.
DR Pfam; PF05986; ADAM spacer1; 1.
DR Pfam; PF01562; Pep M12B propep; 1.
DR Pfam; PF01421; Repolysin; 1.
DR Pfam; PF00090; TSP1; 3.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00209; TSP1; 3.
DR PROSITE; PS0215; ADAM_MPRO; 1.
DR PROSITE; PS00092; TSP1_3.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Integrin.
SQ SEQUENCE 967 AA; 105647 MW; C6349B5D8CBFEA24 CRC64;

Query Match 81.2%; Score 4293; DB 2; Length 967;
• Best Local Similarity 81.2%; Pred. No. 6e-270;
• Matches 791; Conservative 59; Mismatches 110; Indels 14; Gaps 5;

QY 1 MORAVPEGFGRKLGSDMGNAERAPGSRSGFVPVPTLLLLAAA---LLAVSDALGRPSEED 57
DB 1 MOPEVPLSGGLKPKCSMDGDIQRAAKFRSSQSAHMLLLLASITWLLCVRGAHGRPTFED 60
QY 58 BELVVPPELRLVPGHG-TTRLRLHAPDQQLDLVPPDSSFLAPGFTLQNVGRKSGSDPL- 115
DB 61 BELVPLSLERARGHSTTLLRLDAPGQQLHLKLOPDSGLFAPGFTLQTVGRSPGSEAOHL 120
QY 116 -PETDLACHFSGTVNGVPSPSAAALSLCEGVRGAFYLLGLEAFYIQLPA-ASERLATAAP 173
DB 121 DPTGDLACHFSGTVNGVPSPSAAALSLCEGVRGAFYLLGLEAFYIQLPA-ASERLATAAP 180
QY 174 GEKPPAPLQFHLLRRNRQDVGCTGVVDDBPRPTGKAETDEDEDEGTEGEGPQSPQD 233
DB 181 KEESTAPPRPHILRRRRRSGGKGVMDDETLP-----SNSGRESQNTPPQWPLRN 233
QY 234 PALQGVQPTGTSIRKRFVSSHRYVETMLVADQSMABFHGSGLKHVLLTLFSAARLY 293
DB 234 PTPQAGKPTGPGSIRKRFVSSPRYVETMLVADQSMADFHGSGLKHVLLTLFSAARFY 293
QY 294 KHPSTRNSVLSVWVKILVIHDEQGPVETNSAALTNRNFCWQKHNPSPDRDAEHYDTA 353
DB 294 KHPSTRNSVLSVWVKILVIHDEQGPVETNSAALTNRNFCWQKHNPSPDRDAEHYDTA 353
QY 354 ILFTRQDLCSGQTCDTLGMADVGVTCDFSRSCSVIEDDGLQAAFTTAHELGHVFNMFHDD 413
DB 354 ILFTRQDLCSGQTCDTLGMADVGVTCDFSRSCSVIEDDGLQAAFTTAHELGHVFNMFHDD 413
QY 414 AKQASLNGVNDQSHMAASMLSNLDHSPQWPGSCGYMITSPFDNGHCECLMDKQNPQIQL 473
DB 414 AKHCASFNGVSGDHLMAASMLSSLDHSPQWPGSCGYMVTSTFDNGHCECLMDKQNPQIKL 473
QY 474 PGDLPGTGYDANROCOQTFEGSDSKHCPDAASTCTLWCTGTSGGVLCVCKTHPPWADGTS 533
DB 474 PSDLPGLTYDANROCOQTFEGSESTHCPDAASTCTLWCTGTSGGVLCVCKTHPPWADGTS 533
QY 534 CGEGKWCINGKVCVKNRKHKHPDTPFHGSGWMPGWDGCSRTCGGGVQYTWRECDNPVFN 593

DB 534 CGEGKWCVSGKVCVKNKTDKHFATPVHGSWGPWGPWGDGCSRTCGGGVQYTWRECDNPVFN 593
QY 594 GGGKYCEGKRVRYSRNCNLEDPCDNNNGKTFRBEQCAHNEFFSKASFGSGPAVWIPKYAGVS 653
DB 594 GGGKYCEGKRVRYSRNCNLEDPCDNNNGKTFRBEQCAHNEFFSKASFGNEFTVWTPKYAGVS 653
QY 654 PKDRCKLICQAKGIGYFFVLQPKVVDGTPCSPDSTSVCSQCVKAGCDRIIDSKKRPFDK 713
DB 654 PKDRCKLICQAKGIGYFFVLQPKVVDGTPCSPDSTSVCSQCVKAGCDRIIDSKKRPFDK 713
QY 714 CGVCGNGSTCKKISGVSPTSAPGYHDIITPTGATNIEVKQRNQRNNGSFLAKAA 773
DB 714 CGVCGNGSTCKKISGVSPTSAPGYHDIITPTGATNIEVKQRNQRNNGSFLAKAA 773
QY 774 DGTVILNGDYTLSTLEQDIMYKGVVLYRSGSSAALERIRSFSPKEPLTITQVLTVGNALR 833
DB 774 DGTVILNGDYTLSTLEQDIMYKGVVLYRSGSSAALERIRSFSPKEPLTITQVLTVGNALR 833
QY 834 PKIKYTVFVKKKESFNAIPTFSAWIEEWGECSSKCELGWQRRLVECRDINGOPASECA 893
DB 834 PKIKYTVFVKKKESFNAIPTFSEWVIEEWGECSSKCELGWQRRLVECRDINGOPASECA 893
QY 894 KEVPASTRPCADHPCPQWQWGSCKTSGKGYKKTSLKCLSHDGGVLSHSDCDPLKK 953
DB 894 KEVPASTRPCADHPCPQWQWGSCKTSGKGYKKTSLKCLSHDGGVLSHSDCDPLKK 953
QY 954 PKHFIDECTMAECS 967
DB 954 PKHYIDFCILTQCS 967

RESULT 5
ATSL_RAT ID ATSL_RAT STANDARD; PRT; 967 AA.
AC Q9WUQ1; Q9ERI1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE ADAMTS-1 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1).
GN Name=Adamts1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RA Liu X., Tu Y., Yin T., Johnstone E.M., Stephenson D.T., Clemens J.A.,
RA Little S.P.;
RT "Induction of a disintegrin and metalloproteinase with the
RT thrombospondin type I motif (ADAMTS).";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 18-967 FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RA MEDLINE=20304099; PubMed=10847486;
RA Diamantis I., Luehthi M., Hoesli M., Reichen J.;
RT "Cloning of the rat ADAMTS-1 gene and its down regulation in
RT endothelial cells in cirrhotic rats.";
RL Liver 20:165-172 (2000).
CC -I- FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be
CC involved in its turnover. Has angiogenic inhibitor activity (By
CC similarity). Active metalloproteinase, which may be associated with
CC various inflammatory processes as well as development of cancer
CC cachexia. May play a critical role in follicular rupture (By
CC similarity).
CC -I- CATALYTIC ACTIVITY: Cleaves aggrecan at the 1683-Glu-|-Leu-1684
CC site, within the chondroitin sulfate attachment domain.
CC -I- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -I- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).

CC -!- INDUCTION: Down-regulated in endothelial cells derived from
 CC cirrhotic liver.
 CC -!- DOMAIN: The spacer domain and the TSP type-1 domains are important
 CC for a tight interaction with the extracellular matrix.
 CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By
 CC similarity).
 CC -!- SIMILARITY: Belongs to the peptidase M12B family.
 CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
 CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF149118; AAD34012.1; -.
 CC EMBL; AF304446; AAG29823.1; -.
 CC HSSP; P07996; 1LSL.
 CC MEROPS; M12.222; -.
 CC InterPro; IPR006586; ADAM_cysteine.
 CC InterPro; IPR010294; ADAM_spacer1.
 CC InterPro; IPR001762; Disintegrin.
 CC InterPro; IPR006025; Pept_M_Zn_BS.
 CC InterPro; IPR001590; Peptidase_M12B.
 CC InterPro; IPR002870; Peptidase_M12B_N.
 CC InterPro; IPR000884; TSP1.
 CC InterPro; IPR008085; TSP_1.
 CC Pfam; PF05986; ADAM_spacer1; 1.
 CC Pfam; PF01562; Pep_M12B_propep; 1.
 CC Pfam; PF01421; Reprolysin; 1.
 CC Pfam; PF00090; TSP_1; 3.
 CC PRINTS; PR01705; TSP1REPEAT.
 CC SMART; SM00608; ACR; 1.
 CC PROPEP 55 252 By similarity.
 CC CHAIN 233 967 ADAMTS-1.
 CC DOMAIN 253 475 Metalloprotease.
 CC DOMAIN 476 558 Disintegrin-like.
 CC DOMAIN 559 614 TSP type-1 1.
 CC DOMAIN 616 724 Cys-rich.
 CC DOMAIN 725 857 Spacer.
 CC DOMAIN 854 910 TSP type-1 2.
 CC DOMAIN 911 967 TSP type-1 3.
 CC DOMAIN 194 198 Poly-Arg.
 CC SITE 205 205 Cysteine switch (Potential).
 CC METAL 401 401 Zinc (catalytic) (By similarity).
 CC ACT_SITE 402 402 By similarity.
 CC METAL 405 405 Zinc (catalytic) (By similarity).
 CC METAL 411 411 Zinc (catalytic) (By similarity).
 CC CARBOHYD 547 547 N-linked (GlcNAc...) (Potential).
 CC CARBOHYD 720 720 N-linked (GlcNAc...) (Potential).
 CC CARBOHYD 764 764 N-linked (GlcNAc...) (Potential).
 CC CARBOHYD 782 782 N-linked (GlcNAc...) (Potential).
 CC CARBOHYD 945 945 N-linked (GlcNAc...) (Potential).
 CC CONFLICT 21 21 I -> V (in Ref. 2).
 CC CONFLICT 26 31 KFRSQ -> RRGSL (in Ref. 2).
 CC CONFLICT 49 49 R -> A (in Ref. 2).
 CC CONFLICT 72 72 R -> P (in Ref. 2).
 CC CONFLICT 79 79 L -> TR (in Ref. 2).
 CC CONFLICT 249 249 R -> G (in Ref. 2).
 CC CONFLICT 262 265 TMLV -> NLLK (in Ref. 2).
 CC CONFLICT 607 607 S -> F (in Ref. 2).

FT	CONFLICT	936	936	L -> V (in Ref. 2).
FT	CONFLICT	962	962	I -> T (in Ref. 2).
SQ	SEQUENCE	967 AA;	105705 MW;	P93C864F6DCDB4CF CRC64;

Query Match
 Best Local Similarity 81.1%; Score 4287; DB 1; Length 967;
 Matches 790; Conservative 59; Mismatches 111; Indels 14; Gaps 5;

QY	1	MQRAVPEGFGRKLGSDMGNAERAPGSRSPGFVPVPTLLLLAAA--LLAVSDALGRPSSED	57
DB	1	MQPEVPLSGKLPKPCSDMGDIORAAKFFSSQSAHMLLLLLASITMLLCVRGAHGRPTED	60
QY	58	BELVVPPELVRVPGH-GTTRLRHAPDQQLDLDVPPDSFLAPFTLQNVGRKSGSDTFL-	115
DB	61	BELVLPSELRARGHDSITLLRLDAFGQQLHLKLPDPSGLAPFTLTQTVSPGSAQHL	120
QY	116	-PETDLAHCFSYGTGVNGDPSSAAALSICEGVGRGAFYLLGEAFYIQLPLA-ASERLATAAP	173
DB	121	DTGDLAHCFSYGTGVNGDPSSAAALSICEGVGRGAFYLLGEAFYIQLPLA-ASERLATAAP	180
QY	174	GEKPPAPLQFHLRRNRQDVGTCGVVDDEPRPTGKAETEDDEGTEGEGEPQSPQD	233
DB	181	KEESTAPPRFHLRRRRREGSGGAKGVNDETLPT-----SNSGRESQNTPDQWPLRN	233
QY	234	PALQVGOPTGTGTSIRKRFVSSHRYVETMLVAQOSMAEFHSGGLKHVLLTLFVVAARLY	293
DB	234	PTPQAGAKPTGPGSIRKRFVSSHRYVETMLVAQOSMAEFHSGGLKHVLLTLFVVAARFY	293
QY	294	KHPSIRNSVSLVVVKILVIHDEOKPEVTSNAALTNRNFCNWKQHNPPSDRDAEHYDTA	353
DB	294	KHPSIRNSISLVVVKILVIYEEQKPEVTSNAALTNRNFCNWKQHNPPSDRDPHYDTA	353
QY	354	ILFTRQDLCSGOTCDTLGMADVTCDDPSRSCSVIEDDGLQAAFTTAHELGHVFNPMHDD	413
DB	354	ILFTRQDLCSGHTCDTLGMADVTCDDPSRSCSVIEDDGLQAAFTTAHELGHVFNPMHDD	413
QY	414	AKQCASLVNGVNDQSHMMASMLSLNDHSPGSCSYMITSLFDNGHGCLMDKPNPQL	473
DB	414	AKHCASFNGVSGDHLMASMLSLNDHSPGSCSYMITSLFDNGHGCLMDKPNPQL	473
QY	474	PGDLPGTSYDANROCOPTFGEDSKHCPDAASTCTSLWCTGTSGGVLCOTKHPWADGTS	533
DB	474	PSDLPGTLYDANROCOPTFGBESTHCPDAASTCTSLWCTGTSGGVLCOTKHPWADGTS	533
QY	534	CGEGKWCINGKCVNKNHKKHFDTPPHGSGMWGMWPGDCSRCTCGGQVQVYTMRECDNPVKQ	593
DB	534	CGEGKWCVSGKCVNKTDMKHFAIPVHGSWGPWPGDCSRCTCGGQVQVYTMRECDNPVKQ	593
QY	594	GGKYCEGRKRVYRSCNLEDPCDNNNGKTPREOCEAHNEFSKASFGSGPAVEWIPKYAGVS	653
DB	594	GGKYCEGRKRVYRSCNIEDPCDNNNGKTPREOCEAHNEFSKASFGNEPTVEWTPKYAGVS	653
QY	654	PKDRCKLIQAKGIGYFVLOPKVVDGTPCSPDSTSVCGQCVKAGCDRIIDSKKFKDK	713
DB	654	PKDRCKLTCEAKGIGYFVLOPKVVDGTPCSPDSTSVCGQCVKAGCDRIIDSKKFKDK	713
QY	714	GVCGGNGSTCKISGVSVTSKAPGYHDIITPTGATNIEVKORNGSRNNGSFLAIKAA	773
DB	714	GVCGGNGSTCKISGVTSTRPGYHDIITPTGATNIEVKHNRPRGRNNGSFLAIRAA	773
QY	774	DGTYLNGDYLSTLEQDQIMYKVVLYRYSGSSAALERSFSPKLEPTIQLTVGNALR	833
DB	774	DGTYLNGNFTLSTLEQDLTYKGTVLYRYSGSSAALERSFSPKLEPTIQLVMVGHALR	833
QY	834	PKIKTYFVKKKKSFNAIPTFSAWIEWECECSKSELGHQRRLVECRDINGQPASCA	893
DB	834	PKIKTYFPMKKKTFFNPAIPTFSKWIIEWECECSKTSQSGWRVVECRDINGHPASCA	893
QY	894	KEVKPASTRPCADHPCPQWQLGEWSSCSKTCGKYKKTSLKCLSHDGVLSHSDCDPLKK	953
DB	894	KEVKPASTRPCADLPFCPRWQVDWSPCSKTCGKYKKTSLKCLSHDGVLSHSDCDPLKK	953
QY	954	PKHFIDFCTMAECS	967

DR EMBL: AJ315733; CAC86014.1; --
 DR HSP: P07996; ILSL.
 DR MEROPS; M12.025; --
 DR Genew; HGNC:16305; ADAMTS15.
 DR MIM: 607509; --
 DR InterPro; IPR010294; ADAM_spacer1.
 DR InterPro; IPR001762; Disintegrin.
 DR InterPro; IPR001590; Peptidase M12B.
 DR InterPro; IPR002870; Peptidase M12B N.
 DR InterPro; IPR001818; Pept M10A_M12B.
 DR InterPro; IPR006025; Pept_M_Zn_BS.
 DR InterPro; IPR000884; TSPI.
 DR Pfam; PF05986; ADAM_spacer1; 1.
 DR Pfam; PF01562; Pep_M12B_propep; 1.
 DR Pfam; PF01421; Reprolysin; 1.
 DR Pfam; PF00090; TSP 1; 3.
 DR SMART; SM00209; TSP1; 3.
 DR PROSITE; PS50215; ADAM_MEPRO; 1.
 DR PROSITE; PS00346; CYSTEINE_SWITCH; FALSE_NEG.
 DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
 DR PROSITE; PS50214; DISINTEGRIN_2; FALSE_NEG.
 DR PROSITE; PS50092; TSP1; 3.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW Extracellular matrix; Glycoprotein; Hydrolase; Metalloprotease;
 KW Repeat; Signal; Zinc; Zymogen.
 FT SIGNAL 1 17 Potential.
 FT PROPEP 18 212 By similarity.
 FT CHAIN 213 950 ADAMTS-15.
 FT DOMAIN 213 427 Metalloprotease.
 FT DOMAIN 428 515 Disintegrin-like.
 FT DOMAIN 516 571 TSP type-1 1.
 FT DOMAIN 572 700 Cys-rich.
 FT DOMAIN 701 838 Spacer.
 FT DOMAIN 839 895 TSP type-1 2.
 FT DOMAIN 896 949 TSP type-1 3.
 FT SITE 174 174 Cysteine switch (Potential).
 FT METAL 361 361 Zinc (catalytic) (By similarity).
 FT ACT_SITE 362 362 Zinc (catalytic) (By similarity).
 FT METAL 365 365 Zinc (catalytic) (By similarity).
 FT METAL 371 371 Zinc (catalytic) (By similarity).
 FT CARBOHYD 141 141 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 591 591 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 623 623 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 679 679 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 950 AA; 103286 MW; 5DFBE18285CCCC3B CRC64;

Query Match 46.9%; Score 2477.5; DB 1; Length 950;
 Best Local Similarity 48.9%; Pred. No. 4.2e-152;
 Matches 485; Conservative 153; Mismatches 252; Indels 101; Gaps 24;

QY 36 LLLAALAVSDALGRPEDEELVVP--ELE-----RVP---GHGTRRLRHAF 81
 Db 1 MLLGLTLFAPRTAGGSEPEVVPRLDPDINGRRYYWRGDESDGQGLIFQITAF 60
 QY 82 DQQLDLVDPDSSFLAPGFTLQNGRKSGSDTFLP-----ETDLACHFYSGTVNGDPSSA 136
 Db 61 QEDFYHLTPDQAFLAPASTEHLG-----VPLQLTGGSSDLRRCFYSGVDNAEPDSF 114
 QY 137 AALSCEGVGRGAFYLLGEAYFTQPLPAASERLATAAPGKPPAPLQFHLRLRRNQDVG 196
 Db 115 AAVSLCGGLRGAFYRGAEYVISPFPNAS---APAAQRNSQGA---HLLQ---RRGVPG 165
 QY 197 TCGVVDDEPRPTGKAETDEDEGTGEDGPWS-----PQDPALGVQCP--TGTS 247
 Db 166 PSG-----DPTSRC-----GVASG--WNPAIRLALDPYKRRAGFESSRRRS 207
 QY 248 IRKKRFVSHRYVETMLVADQSMABFPHGSLKHYLLTFLPSVAARYLKHPSIRNSVLVV 307
 Db 208 GRAKRFVSIPIRYVELTVVADESVMKFGADLEHYLLTLTATARLYRHSILNPINVV 267
 QY 308 KILVTHDEQKPEVTSNAALTIRNFCNQHNPPSDRDAEYDTAILFTRODLCSQTC 367
 Db 268 KVLRLDRDSGPKVTGNAALTIRNFCNQHNPPSDRDAEYDTAILFTRODLCSQTC 327

RESULT 8
 AT58 HUMAN
 ID AT58 HUMAN STANDARD; PRT; 890 AA.
 AC Q9UP79; Q9NZS0;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE ADAMTS-8 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
 DE with thrombospondin motifs 8) (ADAM-TS 8) (ADAM-TS8) (METH-2) (METH-
 DE 8).
 GN Name=ADAMTS8; Synonyms=METH2;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=99367466; PubMed=10438512; DOI=10.1074/jbc.274.33.23349;
 RA Vazquez F., Hastings G., Ortega M.-A., Lane T.F., Oikemus S.,
 RA Lombardo M., Iruela-Arispe M.L.;
 RT "METH-1, a human ortholog of ADAMTS-1, and METH-2 are members of a new
 RT family of proteins with angio-inhibitory activity.";
 RL J. Biol. Chem. 274:23349-23357(1999).
 RN [2]
 RP SEQUENCE OF 195-440 FROM N.A.
 RX MEDLINE=20079168; PubMed=10610729; DOI=10.1006/geno.1999.6014;

QY 368 DTGLMADYGTCDPSPRSVIEDDGLQAAFTTAHELGHVFNMPHDDAKOCASLNGVNOQS 427
 Db 328 DTGLMADYGTCDPKRSCSVIEDDGLPSAFTTAHELGHVFNMPHNDVVCSEVFGKLRAN 387
 QY 428 HMASMLNLDHSPQWSPSCGYMITSFLDNGHGECIMDKPQNPQIOLPGDLPTSVDANRQ 487
 Db 398 HMSPTLIQIDRANPWSACSAIIITDLDGSHGDCLLDQPSKPISLPDLPGASYTLQQ 447
 QY 498 QFTGEDSKHCPDAASTCSTLWCTGTSGVLVCTQKHPWADGTSCEGKWCINGKCVN 547
 Db 448 CELAFGVGSKPCP-YMQVCTKLWCTGKAKQWVCQTRHFPWADGTSCEGKCLKGACVE 506
 QY 548 K-NHRKHFDTTPHSGWGMWPGDCSRTCCGGVQVYTMRECDNPVPKNGKCYCEGRVYR 606
 Db 507 RHNLNKH--RVDGSWAKWDVPGPCSRCTCGGVQLARRQCTNPTPANGKCYCEGRVYR 563
 QY 607 SCNLEDCPDN-NGKTFREEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPKRCCLKICQAK 665
 Db 564 SCNLEPCPSASGSKSFREEQCEAFNGYHNHSTNRLTLAVAWVPKYSGVSPRDKCKLICRAN 623
 QY 666 GIGYFFVLQPKVVDGTPCSPDSTSVCGOQCVKAGCDRIIDSKKKFKDCGVCGGNGSTCK 725
 Db 624 GTGFYVLAPKVVDTLCSPTSTSVCGVKIKAGCDGNLGSKKRFDKCGVCGGDNKSKCK 683
 QY 726 KISGSVTSKAPGYHDIITPTGATNIEVQRNQRNNGSFLAIKAAADGTIYILNGDYTL 785
 Db 684 KVTGLFTKPMGYNFVVAIPAGASSIDIRQGYKLGIDGDNVLAALKNSQGYLLNGHFV 743
 QY 786 STLEQDINMKGVLRYSGSSAALERIRSFPLKEPTITQVLTVGNALPKIKYTFVKKK 845
 Db 744 SAVERDLVVGSLRLYSQGTAVESLQASRPLEPLTVEVLSVGKMTTPRVRYSPYLPKE 803
 QY 846 ---KES-----FNALPTFS-----AWVIEEWGECSSKCELCGWQRRL 878
 Db 804 PREDKSHPKDPRGSPVJHNSVLSNSVQEPDPPARWVAGSWGPCSCSGSLQKRA 863
 QY 879 VECRDINGQ---PASECAKEVKPASTRCPADHPQWOLGEWSSCKTCGKGYKTKSLKC 935
 Db 864 VDCRGSAGQRTVPACDAAH--RPVETQACGE-PCPTWELSAWSPCKSCSGRFGRRSLKC 920
 QY 936 LSHDGGVLSHSDCDPLKPKHPIDFTWAE 966
 Db 921 VGHGGRLLARDQCNLHRKPQE-LDFCVLRPC 950

D8 AKPCESQLCP 889

RESULT 9
ATSB_MOUSE
ID_ATSB_MOUSE STANDARD; PRT; 905 AA.
AC PS7110;
DT 16-OCT-2001 (Rel. 40, Created)
DT 25-OCT-2001 (Rel. 40, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE ADAMTS-8 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
with thrombospondin motifs 8) (ADAM-TS 8) (ADAM-TS8) (METH-2).
GN Name=Adamts8;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
NC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20079168; PubMed=10610729; DOI=10.1006/geno.1999.6014;
RA Georgiadis K.E., Hirohata S., Seidlin M.F., Apte S.S.;
RT "ADAM-TS8, a novel metalloprotease of the ADAM-TS family located on
mouse chromosome 9 and human chromosome 11.";
RL Genomics 62:312-315(1999).
RC -1- FUNCTION: Has anti-angiogenic properties (By similarity).
CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
matrix (By similarity).
CC -1- TISSUE SPECIFICITY: Expressed specifically in adult lung and heart
and low expression during mouse development.
CC -1- DOMAIN: The spacer domain and the TSP type-1 domains are important
for a tight interaction with the extracellular matrix.
CC -1- PMT: The precursor is cleaved by a furin endopeptidase (By
similarity).
CC -1- SIMILARITY: Belongs to the peptidase M12B family.
CC -1- SIMILARITY: Contains 1 disintegrin-like domains.
CC -1- SIMILARITY: Contains 2 TSP type-1 domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@sib-sib.ch).

DR EMBL; AF175282; AAP25805.1; --
DR HSSP; P07996; IUSL.
DR MEROPS; M12.226; --
DR MGD; MG1:1353468; Adamts8.
DR InterPro; IPRO10294; ADAM_spacer1.
DR InterPro; IPRO01762; DisIntegrin.
DR InterPro; IPRO05025; Pept_M_Zn_BS.
DR InterPro; IPRO01590; Peptidase_M12B.
DR InterPro; IPRO00884; TSPI.
DR InterPro; IPRO08085; TSP1.
DR Pfam; PF05986; ADAM_spacer1.1.
DR Pfam; PF01421; Reprolvain.1.
DR Pfam; PF00090; TSP_1; 2.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00209; TSPI; 2.
DR PROSITE; PS02015; ADAM_MEPR0.1.
DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR PROSITE; PS02014; DISINTEGRIN_2; FALSE_NEG.
DR PROSITE; PS00092; TSPI; 2.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR Extracellular matrix; Glycoprotein; Heparin-binding; Hydrolase;
KW Metalloprotease; Repeat; Signal; Zinc; Zymogen.
FT SIGNAL 1 28 Potential.
FT PROPEP 29 228 By similarity.
FT CHAIN 229 905 ADAMTS-8.
FT DOMAIN 229 452 Metalloprotease.
FT DOMAIN 453 541 Disintegrin-like.


```
Db 37 LLLLLLALLPSARLASPLRBEIEIVFPEKLNGSVLPGSGAPARLLLCRLQAFGETLLLEL 96
QY 90 PPDSSFLAPGFTLQNVGRK----SGSDTLPETDLAHCFYSGTVNGDPSAAALSCEG- 144
Db 97 EQDSGVQVEGLTVQYLQAPPELLGAE---PGT-----VLTGTINGDPSVASLHWDGA 148
QY 145 VNGAFVLLGEAFYIQLPAAASERLATAAAGEKPPAPLQPHLLRRNRQGVGTGCVVDDE 204
Db 149 LLGVLYQVGAELHLOPLEGGTNSA--GGPGA-----HILRRK----- 184
QY 205 PRPTGKAETDEDEGTEGEDEGPQSPDPAALQGVQPTGTGTSIRKKRVSSHRYVETML 264
Db 185 -----SPASGQGPNCVKAP-----LGSPSPRPR-RAKRFASLSRFEVTLV 224
QY 265 VADQSMABPHSGSLKHVLLTLFSAARLYKHPSIRNSVSLVVVKILVHDEQKGPVTSN 324
Db 225 VADDKMAAFHAGAGLKRYLLTVMAAAAKAFKHPSIRNPVSLVTRVLVILGSGEGVQVGS 344
QY 325 AALTURNFQKQNHPPSDRAEHYDTAILFTRODLGSGQTCDTLGMADVGVTCVCDPSRS 384
Db 285 AAQTLSRFCAMQRLNTPEDSGPDHFTAILFTQDLGCVSTCDTLGMADVGVTCVCDPARS 344
QY 385 CSVIEDDGLQAAFTTAHELGHVFNPHDDAKOCASLNG--VNQDSHMMASMLSLNDHSPW 443
Db 405 SPCSFARFTDFLDNGYGHCLLDKPEAPLHPVTFPGKYDADROQCLTFGDSRSHCPQLP 464
QY 504 STCSTLWCTGTSGGVVLCOTKHFPWADGTCGEGKWCINGKCVNKHKKHFDTPPHGSMG 563
Db 465 PPCAALWCSGHNLGHAMCQTKHSPWADGTPCPAQACMGGRCLHWDQIQDFNIPOAGWG 524
QY 564 MWGPWGDCSRTCGGVQVYTMRECPNVPKNGKYGCEGRVRYRSCNLEDCPDNNGKTRE 623
Db 525 PWGPMWDCSRTCGGVQVFSRDCRTPVRNNGKYCEGRTRFSCNTEDCPTGSALTRE 584
QY 624 ECEAHNEFSKASFGSGPA--VWIPKYGVSFKDRCKLICKQAKGIGYRPVLOPKVVDGTP 682
Db 585 EQCAAYNHRTDL-FKSPFQPMQVRYTGAVFQDQCKLTQCARALGYVYVLEPRVVDGTP 643
QY 683 CSPDSTSVCVQGCVKAGCDRIIDSKKFKDKCGVCGNGSGTCKKISGVSITSAPGYHDII 742
Db 644 CSPDSSVVCVQGRCTHAGCDRIIGSKKFKDKCMVCGDGGSGCKSGSPKRYGYNVV 703
QY 743 TTPGTATNIEVQNRQGRNNGSFLAIIKADGTIYLDNGDYTLSTLEQDQIMYKGVV-LRY 801
Db 704 TTPAGATHILVRQQGNPHGRS--IYLALKLPDGSYALNGEYTLMPSPDTDVILFGAVSLRY 761
QY 802 SGSSAALERIRSFSPLEPLTIOVLTVGNALRPKIKYTVFVKKKESFNAIPTPSAWV 859
Db 762 SGATAASELTSGHGGLAQLPTLIQVLVAGNPQDTRULRYSFVPRPTPS--TPRPTQDWL 818

RESULT 11
Q6UWA8 PRELIMINARY; PRT; 837 AA.
AC Q6UWA8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE ADAMTS4.
GN ORFNames=UNQ769;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;
RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,
Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
```

```
RA Eaton D., Foster J., Grimaldi C., Gu Q., Haas P.E., Helden S.,
RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
RA Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,
RA Seshagiri S., Simons L., Singh J., Smith V., Stinson J., Vagts A.,
RA Vandagen R., Watanabe C., Wieand D., Woods K., Xie M.H., Yansura D.,
RA Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,
RA Godowski P.;
RT "The secreted protein discovery initiative (SPDI), a large-scale
RT effort to identify novel human secreted and transmembrane proteins: a
RT bioinformatics assessment.";
RL Genome Res. 13:2265-2270(2003).
DR EMBL; AY368886; AAQ9245.1; -.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR006586; ADAM_cysteine.
DR InterPro; IPR010294; ADAM_spacer1.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR006025; Pept_M_zh_BS.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP 1.
DR Pfam; PF05986; ADAM_spacer1; 1.
DR Pfam; PF01421; Reprolysin; 1.
DR PRINTS; PR01705; TSP1; 1.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS00215; ADAM_MPRO; 1.
DR PROSITE; PS00092; TSP1; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN 1.
SQ SEQUENCE 837 AA; 90166 MW; 5DF9C9ACF67CF8F CRC64;

Query Match 40.0%; Score 2117; DB 2; Length 837;
Best Local Similarity 48.9%; Pred. No. 8.9e-129;
Matches 410; Conservative 132; Mismatches 226; Indels 70; Gaps 18;

QY 36 LLLLLAALLAVSDALGRPSEDEELVPEL---ERVPGHT-TRL--RLHAPDQQLDLVDV 89
Db 37 LLLLLLALLPSARLASPLRBEIEIVFPEKLNGSVLPGSGAPARLLLCRLQAFGETLLLEL 96
QY 90 PPDSSFLAPGFTLQNVGRK----SGSDTLPETDLAHCFYSGTVNGDPSAAALSCEG- 144
Db 97 EQDSGVQVEGLTVQYLQAPPELLGAE---PGT-----VLTGTINGDPSVASLHWDGA 148
QY 145 VNGAFVLLGEAFYIQLPAAASERLATAAAGEKPPAPLQPHLLRRNRQGVGTGCVVDDE 204
Db 149 LLGVLYQVGAELHLOPLEGGTNSA--GGPGA-----HILRRK----- 184
QY 205 PRPTGKAETDEDEGTEGEDEGPQSPDPAALQGVQPTGTGTSIRKKRVSSHRYVETML 264
Db 185 -----SPASGQGPNCVKAP-----LGSPSPRPR-RAKRFASLSRFEVTLV 224
QY 265 VADQSMABPHSGSLKHVLLTLFSAARLYKHPSIRNSVSLVVVKILVHDEQKGPVTSN 324
Db 225 VADDKMAAFHAGAGLKRYLLTVMAAAAKAFKHPSIRNPVSLVTRVLVILGSGEGVQVGS 344
QY 325 AALTURNFQKQNHPPSDRAEHYDTAILFTRODLGSGQTCDTLGMADVGVTCVCDPSRS 384
Db 285 AAQTLSRFCAMQRLNTPEDSGPDHFTAILFTQDLGCVSTCDTLGMADVGVTCVCDPARS 344
QY 385 CSVIEDDGLQAAFTTAHELGHVFNPHDDAKOCASLNG--VNQDSHMMASMLSLNDHSPW 443
Db 405 SPCSFARFTDFLDNGYGHCLLDKPEAPLHPVTFPGKYDADROQCLTFGDSRSHCPQLP 464
QY 504 STCSTLWCTGTSGGVVLCOTKHFPWADGTCGEGKWCINGKCVNKHKKHFDTPPHGSMG 563
Db 465 PPCAALWCSGHNLGHAMCQTKHSPWADGTPCPAQACMGGRCLHWDQIQDFNIPOAGWG 524
QY 564 MWGPWGDCSRTCGGVQVYTMRECPNVPKNGKYGCEGRVRYRSCNLEDCPDNNGKTRE 623
Db 525 PWGPMWDCSRTCGGVQVFSRDCRTPVRNNGKYCEGRTRFSCNTEDCPTGSALTRE 584
```

Db 525 PWGPDGDCSRTCGGVQVQFSSRDCTRPVPRNGKCYCEGRRTRFRSNTEDCPTGSGALTRE 584
 QY 624 EQCEAHNEFSKASFGSGPA-VEWIPKYAGVSPKDRCKLIKQAKGIGYFVLPKPVVDGTP 682
 Db 585 EQCAAYNHRTDL-FKSPFGMDWVRYTGAPDQCKLTQCARALGYVYVLEPRVVDGTP 643
 QY 683 CSPDSTSVCVQCQVKAGCDRIIDSKKPKDKGCVGCGNGSTCKKISGVS TSAKPGYHDI 742
 Db 644 CSPDSSVCVQCRCTHAGCDRIIGSKPKDKCWCVCVGGDGGSGSKSGFRKRYGYNV 703
 QY 743 TIPTGATNIEVKQRNQRNNGSFLATKAADGTYLNGDYTLSTLEODIMYKGVV-LRY 801
 Db 704 TIPAGATHILVRQCGNPGHRS--IYLALKLPGSYALNGEYTLMPSTDVLPAGAVSLRY 761
 QY 802 SGSSAALERIRSFSPKLEPLTIOVLTVGNALRPKIKYTFVKKKESFNAITPFSAW 859
 Db 762 SGATASETLSHGHPLAQPLTLQVLVAGNPQDTRLRYSFVPRPTPS-TPRPTPDWL 818
 RESULT 12
 Q6P4Q8
 ID Q6P4Q8 PRELIMINARY; PRT; 837 AA.
 AC Q6P4Q8;
 DT 05-JUL-2004 (Tremblrel. 27, Created)
 DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
 DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
 DE A disintegrin and metalloproteinase with thrombospondin motifs-4,
 4,
 QN Name=ADAMTS4;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grinstead J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smalius D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC063293.1;
 DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
 DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPRO06586; ADAM cysteine.
 DR InterPro; IPRO10294; ADAM spacer1.
 DR InterPro; IPRO01590; Peptidase M12B.
 DR InterPro; IPRO06025; Pept_M_Zn_BS.
 DR InterPro; IPRO00884; TSP1.
 DR InterPro; IPRO00805; TSP_1.
 DR Pfam; PF05986; ADAM_spacer1; 1.

DR Pfam; PF01421; Reprolysin; 1.
 DR Pfam; PF0090; TSP_1; 1.
 DR PRINTS; PRO1705; TSP1REPEAT.
 DR SMART; SM00608; ACR; 1.
 DR SMART; SM00209; TSP1; 1.
 DR PROSITE; PS02015; ADAM MEPRO; 1.
 DR PROSITE; PS00092; TSP1; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
 KW Integrin.
 SQ SEQUENCE 837 AA; 90149 MW; 3B91C651E54EFC5F CRC64;
 Query Match 39.8%; Score 2106; DB 2; Length 837;
 Best Local Similarity 48.7%; Pred. No. 4.6e-128;
 Matches 408; Conservative 134; Mismatches 226; Indels 70; Gaps 18;
 QY 36 LLLLAALLAVDADALGRPSDEBELVPEL---BRVPGHGT-TRL--RLHAFDQOLDLDV 89
 Db 37 LLLLLLASLPASRLASPLPREEBIVFPEKLNGSVLPQSGAPARLLCRLQAFGETLLIEL 96
 QY 90 PPDSSFLAPGFTLONVGRK---SGSDTPLPETDLAHCIFYSGTVNGDPSSAAALSCEG- 144
 Db 97 EQDSGVQVEGLTVQYLQAPELGGAB---PGT-----YLTGTNGDPESVASLHWGGGA 148
 QY 145 VRGAFYLLGEAYFTIPLPAASERLATAPGKPPAPLQFHLRRNRQDVGCTCGVVDDE 204
 Db 149 LLGVLYRGAEHLQPLEGGTPNSA-GPGCA-----HILRK----- 184
 QY 205 PRPTGKAETDEDEGTGEGDEGPQWSPQDPALQGVQPTGTSIRKRRFVSSHRYVETML 264
 Db 185 -----SPASGGQPMCNVKAP-----LGSPSPRPR-RAKRFASLSRFVETLV 224
 QY 265 VADQSMAEFHSGGLKHYLLTLFSVAARLYKHPSIRNSVLSLVVVKILVIHDEQKPEVTSN 324
 Db 225 VADKMAAFHGAGLKYLLTVMAAAAKAFKHPSTRNPVSLVTVTRILVILGSGEGPQVGPS 284
 QY 325 AALTLRNFQKQHPNPPDRDAEHYDTAILFTRQDLCSQTCDTLGMADVGTGCDPSRS 384
 Db 285 AAQTLRSFCAWQRCGLNTPENSDPDHPDTAILFTRQDLCGVSTCDTLGMADVGTGCDPARS 344
 QY 385 CSVIEDGLQAAFTTAHELGHVFNMPHDDAKQACASLNG-VNDSHMMASMLNLDHSGPW 443
 Db 345 CAIVEDGLQSAFTAAHELGHVFNVLHNSKPCISLNGPLSTSRVHMAPVMAHVDPEPW 404
 QY 444 SPGSGYMITSLDNGHCEGLMDKPNQPIQLPGDLPGTSYDANROCOFTFCEDSKHCDDAA 503
 Db 405 SPGARFITDNLNGYGHCLLDKPEAPLHPVTFPGKDYDADROCOQLTFGDSRHCQLP 464
 QY 504 STCSLWCTGTSGGLVYCQTKHFPWADGTSCEGKWCINGKCNKHNKHKHFDTPFHGSWG 563
 Db 465 PPCAALWCSHLNGHAMCQTKHSPWADGTPCGPAQACMGGRCLHMDQLQDFNIPAGGWG 524
 QY 564 MWGPWGDCSRTCGGVQVYTWRECDNPVRNGGKYCEGKRVYRNSCNLEDPCDNNKGTFR 623
 Db 525 PWGPWGDCSRTCGGVQVQFSSRDCTRPVTRNGGKYCEGRRARFRSCNTEDECTGSAITRE 584
 QY 624 EQCEAHNEFSKASFGSGPA-VEMIPKYAGVSPKDRCKLIKQAKGIGYFVLPKPVVDGTP 682
 Db 585 EQCAAYNHRTDL-FKSPFGMDWVRYTGAPDQCKLTQCARALGYVYVLEPRVVDGTP 643
 QY 683 CSPDSTSVCVQCQVKAGCDRIIDSKKPKDKGCVGCGNGSTCKKISGVS TSAKPGYHDI 742
 Db 644 CSPDSSVCVQCRCTHAGCDRIIGSKPKDKCWCVCVGGDGGSGSKSGFRKRYGYNV 703
 QY 743 TIPTGATNIEVKQRNQRNNGSFLAIKAADGTYLNGDYTLSTLEODIMYKGVV-LRY 801
 Db 704 TIPAGATHILVRQCGNPGHRS--IYLALKLPGSYALNGEYTLMPSTDVLPAGAVSLRY 761
 QY 802 SGSSAALERIRSFSPKLEPLTIOVLTVGNALRPKIKYTFVKKKESFNAITPFSAW 859
 Db 762 SGATASETLSHGHPLAQPLTLQVLVAGNPQDTRLRYSFVPRPTPS-TPRPTPDWL 818
 RESULT 13

Qy	130	NGDPSSAALSCEG-VRGAFYLLGEAYFTQPLPAASERLATAAPGEKPPAPLPQFHLLR	181
Db	129	NGDPSSVASHLWDGGLLQVLYQYGAELHLQPLEGGALNSA-GPGGA-----HILRR	179
Qy	189	NRQGDVGGTCGVDDPRPTKAETDEDETEGEDEGEQPOWSQDPAQLQGVGPTGTGSI	248
Db	180	K-----SPASSOGPMCTVKARF-----GSFPI-SR	204
Qy	249	RKRKRFVSHRYVETMLVADQSMABFHGSLGHKHYLLTLFSVAARLYKHPISIRNSVSLVVK	308
Db	205	RTKRFAISLSRFVETLLVAVDDKMAAFHGTGLKRYLLTVMAAAAKAFKHPISIRNPVLVWTR	264
Qy	309	ILVIDHEQKQPEVTSNAALTIRNCFNWQKHNPSPDRDAEHYDTAILFTQDILCGSQTC	368
Db	265	LVLILSGEGEPQVGSAAQTILRSFCTWQRGINTENDSDPDHFDPTAILFTQDILCGSVTCD	324
Qy	369	TLGMADVTVCDPSSSCSVIEDDGLQAFTTAHBLGHVFNMPHDDAKQKASLNGVNQDS-	427
Db	325	TLGMADVTVCDPARSCAIVDDGLQSAFTHAHLGHVFNMLHDNSKPCPTNLNGQGSRR	384
Qy	428	HMWASMLNLDHSQPSWPCSGYMITSLFDLNGHSECLMDKPNQPIQLPGDLPGTYSYDNRQ	487
Db	385	HVMAPVMAHVDPEEPWSPCSARFITDFLDNGYGHCLLDKPEAPLHLPATFPKGXDYADRQ	444
Qy	488	CQFTFGBDSKHCPDAASTCTSLWTGTSGGVLCVQTKHFPWADGTSCEGKWCINKCVN	547
Db	445	COLTFGPSSSHCPQLPPPCAAALWCSGLHNGHAMQTKHSPWADGTPCGSSQACMGGRCLH	504
Qy	548	KXHKRHGFDTPFHSGWMGPMGDCSRCTCGGQVOVTMRECDNVPKNGKGYCEGKRVYRS	607
Db	505	VQDLKDFNVQAGWPMGPMGDCSRCTCGGQGVQSSRDCTRPVPRNGGKYCEGRRTRFRS	564
Qy	608	CNLEDCPDNNGKTFREEOCEAHNEFASKASFGSPA-VEMIPKYAGVSPKDRCKLIQAKG	666
Db	565	CNTENCPHGSALTFREEOCAAYNHRD-L-FKSPPGMDWVPRTGTGVAPRDOCKLTQCARA	623
Qy	667	IGYFVFLQPKVVDGTPCSPDSTSVCGVQCVKAGCDRIIDSKKKFKDKGVCGNGSTCKK	726
Db	624	LGYYVVLPRVADGTPCSPDSTSVCGVQGRCTHAGCDRIIGSKKKFKDKMVCVCGDGRCSK	683
Qy	727	IGSVTSAPKGYHDIITPTGATNIEVKQRNQRSGRNSGFLAKAADGTVILLNGDYTILS	786
Db	684	QSGSFKKFRYGSVDVVTIPAGATHILVRQ--QGSGSLKSYLLALKUSDGSAUNGEBYTLM	741
Qy	787	TLEQDIMYKGVV-LRYSGSSAALERISFSPKLEPTIQVLTVGNALRPKIKYTFVVKK	844
Db	742	PFETDVVLFGAVLSRYSGATASETLSGHGFLAQPLTLQVLVAGNQAPLARLSVFFVPR	800
RESULT	14		
Q8BNJ2	ID	Q8BNJ2	PRELIMINARY; PRT; 845 AA.
AC	Q8BNJ2;		
DT	01-MAR-2003	(TrEMBLrel. 23, Created)	
DT	01-MAR-2003	(TrEMBLrel. 23, Last sequence update)	
DE	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)	
DE	Mus musculus	9 days embryo whole body cDNA, RIKEN full-length enriched library, clone.D030041M02 product:a disintegrin-like and metalloprotease (reprolysin type) with thrombospondin type 1 motif, 4, full insert sequence.	
DE	NCBI	Taxid=10090;	
GN	Name=Adamts4;		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_Taxid=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=Whole body;		
RX	MEDLINE=9297253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;		
RA	Carninci P.; Hayaehizaki Y.;		
RT	"High-efficiency full-length cDNA cloning.";		
RL	Meth. Enzymol. 303:19-44(1999).		

[2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RX MEDLINE=2108560; PubMed=11217851; DOI=10.1038/35055500;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 [3]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 [4]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.,
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 [5]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RA "RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 [6]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RA "RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 [6]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akaira S., Takeda Y., Tanaka T.,
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK083534; BAC38944.1; -.
 DR HSSP; P07996; 1LSL.
 DR MEOPS; M12.221; -.
 DR MGD; MGI:1339849; Adante4.
 DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
 DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR010294; ADAM spacer1.
 DR InterPro; IPR001590; Peptidase M12B.
 DR InterPro; IPR006025; Pept_M_Zn_BS.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR008085; TSP 1.
 DR Pfam; PF05986; ADAM_spacer1; 1.
 DR Pfam; PF01421; ReprGlysin; 1.
 DR Pfam; PF00090; TSP 1; 1.
 DR PRINTS; PR01705; TSP1REPEAT.
 DR SMART; SM00209; TSP1; 1.
 DR PROSITE; PS00215; ADAM_MPRO; 1.
 DR PROSITE; PS00092; TSP1; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.

KW Integrin; Metalloprotease; Protease.
 SQ SEQUENCE 845 AA; 91238 MW; B496C3190D1A9225 CRC64;
 Query Match 39.8%; Score 2104.5; DB 2; Length 845;
 Best Local Similarity 49.2%; Pred. No. 5.9e-128;
 Matches 405; Conservative 131; Mismatches 216; Indels 71; Gaps 18;
 QY 37 LLLAAALLAVSDALGRPSEDEELVPE---LERVPGHGT-TRL--RLHAFDQDLDDV 89
 DB 46 LLLAFLLSLAMP-ASPLPREEEIVFPEKLNGSSILFOSGVPARLLYRLPAPFAGEMLLLEL 104
 QY 90 PPDSSFLAPGFTLQNVGRK-----SGSDTPLPETDLAHCFTYSGTVNGDPSSAAALSCLSG- 144
 DB 105 EQDPGVQVEGLTVQYLGQAPMLGAB---PGT-----YLTGTINGDPESVASLHWDGGA 156
 QY 145 VRGAFYLLGEAYFTIQLPAASERLATAAPGKPPAPLOFHLRLRRNRQGDVGCTGCVVDDE 204
 DB 157 LLGLVYRGAEHLQLPLEGGALNSA-GPGGA-----HILRRK----- 192
 QY 205 PRPTGKAETDEDESGTEGDEGPQWSPQDPALQGVGPTGTGSTRKKRFVSSHRVYETML 264
 DB 193 -----SPASSQGPMTVKAPS-----GSPSPI-SRTRKRFASLSRFVETLV 232
 QY 265 VADQSMAEFFHGSLKHYLLTLFSVAARLYKHPISVLSVVKILVIHDEKQGPVTSN 324
 DB 233 VADDKMAAFHGTGLKRYLLTVMAAAAKAFKHPISIRNPVNLVTVLVLGSGQEGPQVGPS 292
 QY 325 AALTLRFNCWQKOHNPSPSDAEHYDTALTFTQDLCCGSDCTCLGMADVGVTCDFSR 384
 DB 293 AAOTLRSFCWQGLNTPNSDPDFDTALTFTQDLCCGSDCTCLGMADVGVTCDFPAR 352
 QY 385 CSVIEDDGLQAAFTTAHELGHVFNHPHDDAKOCASLNGVNODS-HMVASMLSLNDHSDPW 443
 DB 353 CAIVEDDGLASAFTHAHELGHVFNHLDNKSPTNLNGQSSSRHVWAPVMAHVDPEPW 412
 QY 444 SPCSGYMITFSLDNGHCECLMDKPNPIQLPGDLPGTSYDANRQCQFTFGEDSKHCPDAA 503
 DB 413 SPCSARFITDFLDNGYGHCLLDKPEAPLHLPATPPGKDYADADRCQLTFFGPDSSHCPLP 472
 QY 504 STCSTLACTGTSGVLVCOIKHFPWADGTSCGEGKWCINGKVCYNKHKHFDTPPHGSMG 563
 DB 473 PPCAALWCSGHLNHAMCQTKHSPWADGTPCGSSQACMGGRCLHVDQLKDFNVPOAGWG 532
 QY 564 MWGPWGDCSRTCGGVQVWRECDNPNVPKNGKCEGKRVRYRSCNLEDCPDNNGKTFRE 623
 DB 533 PWGPWGDCSRTCGGVQVFSRDDCTRPVPRNGKCEGRTFRSCNTENCCHGSALTFR 592
 QY 624 EQCEAHNEFSKASFGSGPA-VEMIPKYAGVSPKDRCKLIQAKGIGYFFVLQPKVVDGTP 682
 DB 593 EQCAAYNHRDGL-FKSPFGPMWVPRVTGVAPRDQCLTQCARALGYVYVLEPRVADGTP 651
 QY 683 CSPDSTSVCCVQCQKACDRIIDSKKFKDKCGVCGGNGSTCKKISGVSITSAKPGYHDI 742
 DB 652 CSPDTSVCCVQCQKACDRIIGSKKKFKDKMWCGGSDGRCQSGSKQSGFKFRYGVSDV 711
 QY 743 TIPTGATNIEVKQRNQRSGRNNGSFLAIKAADGTYILNGDYTLSTLEQDIMYKGVW-LRY 801
 DB 712 TIPAGATHILVRQ--QGGSLGKSIYLAALKLSDGSYALNGEYTLMPSTPDVVLPGAVSLRY 769
 QY 802 SGSSAALERIRSFPLKEPTIQLVTGNALRPKIKTYTFVKK 844
 DB 770 SGATAASETLHGHLAQLTLQVLVAGNPNQARLRYSFVPR 812
 RESULT 15
 QYYS95 PRELIMINARY; PRT; 839 AA.
 ID QYYS95
 AC QYYS95;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE AggreCanase-1.
 GN Name=ADAMTS-4;

This Page Blank (uspto)